

CC A bifunctional protein (AA85505) consists of a single chain antibody, scFv(FRP)5, directed against the tumour erbB-2 antigen, a hinge region, and a functional zeta chain obtd. from a T-cell receptor. The protein is expressed in host cells, esp. cytotoxic T-lymphocytes, providing them with a defined tumour cell specificity enabling targeting to defined tumour cells and MHC-unrestricted and MHC-independent tumour destruction in vitro or in vivo

XX Sequence 461 AA;

Query Match 100.0%; Score 2441; DB 2; Length 461;
Best Local Similarity 100.0%; Pred. No. 1.4e-160;
Matches 461; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAMWTLFLMAAKPKQIQIVOSGPBLKKPEYTKISCKASGYFTYGMNWKQADG 60
DB 1 MAMWTLFLMAAKPKQIQIVOSGPBLKKPEYTKISCKASGYFTYGMNWKQADG 60
QY 61 OGKMMGWINTSGESTFADDFKGRFDSLETSANTAYQIINNLSKEDMATYFCARMEY 120
DB 61 OGKMMGWINTSGESTFADDFKGRFDSLETSANTAYQIINNLSKEDMATYFCARMEY 120
QY 121 HGVPYWGQTTVYVSSGGSGSGGSDIQLTQSHKFLSTSVGDRVSITCKASQD 180
DB 121 HGVPYWGQTTVYVSSGGSGSGGSDIQLTQSHKFLSTSVGDRVSITCKASQD 180
QY 181 VYNAVAYWQKPGQSPKLLIYASSTRYTPSPRFTGSGGSPPTFTISSVQMEDLAIVFC 240
DB 181 VYNAVAYWQKPGQSPKLLIYASSTRYTPSPRFTGSGGSPPTFTISSVQMEDLAIVFC 240
QY 241 OQHFRTPFGSGTKLEIKALEISNSVMYFSSVVPVLQKVNSTTKPVLRTPSPVHPGT 300
DB 241 OQHFRTPFGSGTKLEIKALEISNSVMYFSSVVPVLQKVNSTTKPVLRTPSPVHPGT 300
QY 301 SOPQRPEDCRPRGSYKGTGLDFLEDPKLCYLLDGLFTYGVITLALYRAKFSRSAAETPA 360
DB 301 SOPQRPEDCRPRGSYKGTGLDFLEDPKLCYLLDGLFTYGVITLALYRAKFSRSAAETPA 360
QY 361 NLQDPNQLYNELNGRREYDYLEKKRARDPEMGKQORRRRPPQGVNNAQKDMAEAY 420
DB 361 NLQDPNQLYNELNGRREYDYLEKKRARDPEMGKQORRRRPPQGVNNAQKDMAEAY 420
QY 421 SEITGERRRGGKHDGLYQGLSTATKDTYDALHMOTLAPR 461
DB 421 SEITGERRRGGKHDGLYQGLSTATKDTYDALHMOTLAPR 461

RESULT 2

AAW82315
ID AAW82315 standard; protein; 483 AA.

XX AAW82315;

XX 26-FEB-1999 (first entry)

XX Chimeric CD19/cTCR protein.

XX Single chain; antibody; B-cell marker; CD19; scFvCD19; gene therapy;
XX retroviral vector; tumour-associated antigen; cancer; immunoglobulin;
XX CDS alpha chain; hinge region; theta chain; transmembrane domain;
XX intracellular domain; T-cell receptor-CD3 complex.

XX Homo sapiens.

XX Synthetic.

XX DE19720152-A1.

XX 05-NOV-1998.

XX 02-MAY-1997; 97DE-01020152.

XX 02-MAY-1997; 97DE-01020152.

PA (DELB-) DELBRUECK CENT MOLEKULARE MEDIZIN MAX.

XX Schroeder A, Schwenkenbecher J, Bargou R;

XX WPI; 1998-584686/50.

XX DR N-PSDB; AAW73324.

XX Retroviral vector system for gene therapy - by expression of therapeutic gene in T cells.

XX Claim 9; Page 6; 10pp; German.

CC This sequence represents a novel chimeric CD19/cTCR protein which is composed of an immunoglobulin leader sequence, an anti-CD19 single chain antibody fragment (scFvCD19), a hinge region from the CD8 alpha-chain and the transmembrane and intracellular domain of the T-cell receptor-CD3 complex theta chain. This sequence is used in a retroviral vector system for gene transfer and expression in primary human T cells which involves a retroviral vector containing cDNA corresponding to a therapeutic gene, a promoter selected from the SV40 immediate early promoter or the human IL-2 or MHC I promoter and a selectable marker, and a packaging cell line (i.e. a modified retroviral packaging line for enhanced transfer into CC primary human T lymphocytes). This vector system in which the cDNA encodes a single-chain antibody (scFv) specific for a tumour-associated antigen can be used for gene therapy of cancer, especially by isolating T cells from the patient's blood, stimulating the T cells in vitro, introducing the retroviral vector into the T cells and returning the T cells to the patient, preferably by intravenous injection

XX Sequence 483 AA;

Query Match 68.6%; Score 1675; DB 2; Length 483;
Best Local Similarity 69.2%; Pred. No. 1.4e-107;
Matches 334; Conservative 41; Mismatches 86; Indels 22; Gaps 6;

QY 1 MAMWTLFLMAAK-VPKQIQIVOSGPBLKKPEYTKISCKASGYFTYGMNWKQADP 59
DB 1 MDWIMRILFLVGAATGASHSEVQLOESGALVRRPSSVKISCKASGYAFSSYGMNWKQADP 60
QY 60 OGKMMGWINTSGESTFADDFKGRFDSLETSANTAYQIINNLSKEDMATYFCARMEY 119
DB 61 OGKMMGWINTSGESTFADDFKGRFDSLETSANTAYQIINNLSKEDMATYFCARMEY 120
QY 120 ----YHGVYWGQTTVYVSSGGSGS--GGSGSGSGSDIQLTQSHKFLSTSVGDRVS 172
DB 121 TTGVPYVANDYWGQTSCTAASAKTTPKLEBGFSEARVDILITQSPASLAVALGQAT 180
QY 173 ITCRASQDV-YNA---VAMYQOKPGQSPKLLIYASSTRYTPSPRFTGSGGSPPTFTIS 228
DB 181 ISCRASQSDYDGDYSVLNMYQQLPGQPKLLIYASSTRYTPSPRFTGSGGSPPTFTIS 240
QY 229 SVQMEDLAIVFCQOHPFTPTFGSGTKLEIKALEISNSVMYFSSVVPVLQ 278
DB 241 PVEKVDATATYHCOOSTEDPWTFGGTGLTEILEOKLISEEDLNLEISNSVMYFSSVVPVLQ 300
QY 279 KVNSTTKPVLRTPSPVHPGTSPORPREDCRPRGSYKGTGLDFLEDPKLCYLLDGLFT 338
DB 301 KVNSTTKPVLRTPSPVHPGTSPORPREDCRPRGSYKGTGLDFLEDPKLCYLLDGLFT 360
QY 339 YGVITLALYRAKFSRSAAETPAANLQDPNQLYNELNGRREYDYLEKKRARDPEMGKQ 398
DB 361 YGVITLALYRAKFSRSAAETPAANLQDPNQLYNELNGRREYDYLEKKRARDPEMGKQ 420
QY 399 RRRNPOGVNNAQKDMAEAYSEITGERRRGGKHDGLYQGLSTATKDTYDALHMOTL 458
DB 421 RRRNPOGVNNAQKDMAEAYSEITGERRRGGKHDGLYQGLSTATKDTYDALHMOTL 480
QY 459 APR 461
DB 481 APR 483

RESULT 3

[illegible]

DB		399	GKGHDLGYOGLSTATCTDYLALHMQTLAPR	428
	RESULT 5			
ID	AAR95056			
XX	AAR95056 standard; protein; 615 AA.			
AC	AAR95056;			
DT	16-OCT-2003 (revised)			
DT	19-AUG-1996 (first entry)			
XX				
DE	GAL4-DT-scfv(FRP5) multidomain protein.			
XX				
KM	Nucleic acid transfer system; gene transfer; gene therapy;			
KW	cell targeting; multidomain protein; vector; cancer; GAL4;			
XX	diphtheria toxin; single chain antibody; scfv; FRP5.			
OS	Saccharomycetes; cerevisiae.			
OS	Clostridium; diphtheriae.			
OS	Mus sp.			
OS	Chimeric.			
OS	Synthetic.			
FH	Key	Location/Qualifiers		
FT	Peptide	1..8		
FT	/label= FLAG_epitope			
FT	Peptide	9..15		
FT	/label= Spacer			
FT	Domain	16..161		
FT	/label= GAL4			
FT	/note= "amino acids 2-147 of yeast GAL4"			
FT	Peptide	162..165		
FT	/label= Spacer			
FT	Domain	166..354		
FT	/label= DT			
FT	/note= "amino acids 156-384 of diphtheria toxin"			
FT	Peptide	355..361		
FT	/label= Spacer			
FT	Domain	362..601		
FT	/label= scfv(FRP5)			
FT	Peptide	602..615		
FT	/label= Spacer			
PN	WO9613599-A1.			
PN				
PD	09-MAY-1996.			
PF	31-OCT-1995; 95MO-EP04270.			
PR	01-NOV-1994; 94EP-00810627.			
PA	(WELS/) WELS W.			
PI	Weiss W, Fominaya J;			
XX	WPI; 1996-239505/24.			
DR	N-PSDB; AAT29412.			
XX				
PT	Nucleic acid transfer system for gene therapy, e.g. against cancer - includes toxin translocation domain to target nucleic acid to specific cell.			
PT				
XX				
PS	Claim 7; Page 81-83; 106pp; English.			
CC	A multidomain protein (AAR95056) has a FLAG epitope, the DNA binding domain of yeast GAL4, a portion of diphtheria toxin that acts as a translocation domain, and a single chain antibody, scfv, of monoclonal antibody FRP5 (raised against human tumour cell HBR2 antigen) that act as a ligand domain. It is the product of a fusion gene (AAT29412), and used with an effector nucleic acid that comprises e.g. a gene to be delivered to a cell and a cognate structure for the GAL4 DNA binding domain. This provides a novel means of nucleic acid transfer, suitable			

CC		for gene therapy.	(Updated on 16-OCT-2003 to standardise OS field)
XX	Sequence	615 AA;	
SQ	Query Match	53.6%; Score 1308; DB 2; Length 615;	
	Best Local Similarity	96.5%; Pred. No. 4.6e-82;	
	Matches	245; Conservative 3; Mismatches 6; Indels 0; Gaps 0	
OY	19 QIQLVQSPELKKPGETVKISCKASGYPTNYGMNVKQAPQGLKMMGMINSTGESEF	78	
DB	362 QVLQQSQSPELKPKPGETVKISCASAGYPPTYNGMNVKQAPOGLKMMGINSTGESEF	421	
OY	79 ADDFRGRDPSLETSANTANAYLOINNLKSDDATTCARREYHGVPPYWGOSTTYTVSSG	138	
DB	422 ADDFRGRDPDSLETSANTANAYLOINNLKSDMATYCARREVHGVPPYGOGTTVTVSIG	481	
OY	139 GGSGCGGGGGGGSPDIQTQSHKFPLSTSVDGIVSLTCKASODPVNNAVAVYOOKPGOSPFL	198	
DB	482 GGSGCGGGGGGGGGSPDIQTQSHKFPLSTSVDGIVSLTCKASODPVNNAVAVYOOKPGOSPFL	541	
OY	199 LIYSASSRYTGVPNRFTSGSGPDFTTISSVOAEDLAVALFCQHFRTPFTFGSGTKLEI	258	
DB	542 LIYSASSRYTGVPNSRFSTGSGGPDFTTISSVOAEDLAVALFCQHFRTPFTFGSGTKLEI	601	
OY	259 KALEINSNMVFSS 272 : : : :		
DB	602 KALSDLSSRRFFSA 615		
<hr/>			
RESULT 6			
AAR95057	ID AAR95057	standard; protein; 617 AA.	
XX	AA95057;		
DT	27-AUG-2003	(revised)	
DT	19-AUG-1996	(first entry)	
XX	GALA-DT-scfv(FRP5) multidomain protein.		
XX	Nucleic acid transfer system; gene transfer; gene therapy;		
KW	cell targeting; multidomain protein; vector; cancer; GAL4;		
KW	diphtheria toxin; single chain antibody; scfv; FRP5.		
XX	Saccharomycetes; cerevisiae.		
OS	Corynebacterium diphtheriae.		
OS	Myx sp.		
OS	Chimeric.		
OS	Synthetic.		
XX	Key	Location/Qualifiers	
FT	Peptide	1	
FT	Peptide	/label= Spacer	
FT	Peptide	2..9	
FT	Peptide	/label= FLAG_epitope	
FT	Peptide	10..17	
FT	Domain	/label= Spacer	
FT	Domain	18..163	
FT	Peptide	/label= GAL4	
FT	Peptide	/note= "amino acids 2-147 of yeast GAL4"	
FT	Peptide	164..167	
FT	Domain	/label= Spacer	
FT	Domain	168..356	
FT	Peptide	/label= DT	
FT	Peptide	/note= "amino acids 196-384 of diphtheria toxin"	
FT	Peptide	357..363	
FT	Domain	/label= Spacer	
FT	Domain	364..603	
FT	Peptide	/label= scfv(FRP5)	
FT	Peptide	604..617	
XX	Peptide	/label= Spacer	
XN		MO9613599-Al	

Query Match	53.6%	Score 1308	DB 2	Length 617
Best Local Similarity	96.5%	Pred. No. 4,6e-82		
Matches 245	Conservative 3	Mismatches 6	Indels 0	Gaps 0
09-MAY-1996.				
31-OCT-1995;				
95WO-EP004270.				
01-NOV-1994;				
94EP-00810627.				
(WELLS/) WELLS W.				
Wells W, Fominaya J;				
WPI, 1996-239505/24.				
N-P8DB; AAT29413.				
Nucleic acid transfer system for gene therapy, e.g. against cancer - includes toxin translocation domain to target nucleic acid to specific cell.				
Claim 7, Page 87-89, 106pp; English.				
A multidomain protein (AAR95057) comprises a FLAG epitope, the DNA binding domain of yeast Gal4, a portion of diphtheria toxin that acts as a translocation domain, and a single chain antibody, scFv, of monoclonal antibody FRP5 (raised against human tumour cell HER2 antigen) that acts as a ligand domain. It is the product of a fusion gene (AAT29413), and is used with an effector nucleic acid that comprises e.g. a gene to be delivered to a cell and a cognate structure for the Gal4 DNA binding domain. This provides a novel means of nucleic acid transfer, suitable for gene therapy. (Updated on 27-AUG-2003 to correct OS field.)				
Sequence 617 AA:				
Query Match	53.6%	Score 1308	DB 2	Length 617
Best Local Similarity	96.5%	Pred. No. 4,6e-82		
Matches 245	Conservative 3	Mismatches 6	Indels 0	Gaps 0
19 QIQLVQSGPELKKPGETVKISCKASGYPTNYGMWVQAPOGQKMKMGINTSTGESTF 78				
364 QVQLQQSGPELKKPGETVKISCKASGYPTNYGMWVQAPOGQKMKMGINTSTGESTF 423				
79 ADDPFGKRPDSLETSSANTAYVQIINLKSEDMATFCARMEVYHGVVPYQGQGTTVVSSG 138				
424 ADDPFGKRPDSLETSSANTAYVQIINLKSEDMATFCARMEVYHGVVPYQGQGTTVVSSG 483				
139 GCGSGGSGGSGGSDIQLTQSHKFLSTVGRVSIITCKASQDVYNAVAVYQKPGQSPKL 198				
484 GCGSGGSGGSGGSDIQLTQSHKFLSTVGRVSIITCKASQDVYNAVAVYQKPGQSPKL 543				
199 LIYSASSRYTGVPSHFTSGSGGPDFTFTISSVQAEIDLAVYFCQGHFRPTFTGSGTKLEI 258				
544 LIYSASSRYTGVPSHFTSGSGGPDFTFTISSVQAEIDLAVYFCQGHFRPTFTGSGTKLEI 603				
259 KALEISNSVMTFSS 272				
604 KALELSSERRFSA 617				
RESULT 7				
AAR26982				
AAR26982 standard; protein, 637 AA.				
AAR26982;				
24-OCT-2003 (revised)				
25-MAR-2003 (revised)				
11-FEB-1993 (first entry)				
(FRP5) -ETA fusion protein.				
Monoclonal antibody; light chain; heavy chain; tumour; c-erbB-2; variable region, ETA.				
Pseudomonas aeruginosa; PAK.				

```

XX Key Location/Qualifiers
FH Peptide 1..21
FT /label= ompa_signal_peptide
FT Peptide 22..29
FT /label= FLAG_peptide_and_enterokinase_cleavage_site
FT Domain 33..151
FT /label= FRP5_heavy_chain_variable_domain
FT Peptide 152..166
FT /label= Linker
FT Domain 167..274
FT /label= FRP5_light_chain_variable_domain
FT Protein 276..397
FT /label= ETA_252-613
XX EP502812-A1.
XX 09-SEP-1992.
XX 27-JAN-1992; 92EP-00810056.
XX 05-FEB-1991; 91EP-00810079.
XX (CIBA ) CIBA GEIGY AG.
XX Wels WS, Hynes NE, Harwerth I, Groner B, Hardman N, Zwickl M;
XX WPI: 1992-302096/37.
XX N-PSDB; AAQ28257.
XX Recombinant antibodies directed to growth factor receptor C-ERBB-2 - for
XX diagnosing and treating tumours expressing C-ERBB-2 e.g. breast or
XX ovarian tumours.
XX PS Disclosure: Page 47-52; 67pp; English.
XX The sequences given in AAR26982-3 contain part of the exotoxin A (ETA)
XX sequence corresponding to positions 252-613 of the full exotoxin A
XX sequence. These sequences are encoded by Fv(FRP5)-ETA fusion genes. The
XX ETA sequence was used as a marker gene so that E. coli transformed with
XX the fusion gene could be identified. The fusion genes were expressed in
XX E. coli and the antibodies were extracted. These recombinant antibodies
XX can be used for the qualitative and quantitative determination of c-erbB-
XX 2. This can be used for monitoring or in-vivo localisation of tumours
XX overexpressing c-erbB-2. (Updated on 25-MAR-2003 to correct PN field.)
XX CC (Updated on 24-OCT-2003 to standardise OS field)
XX CC
XX CC
XX Sequence 637 AA;

```

```

Query Match 53.4%; Score 1304; DB 2; Length 637;
Best Local Similarity 97.6%; Pred. No. 9e-82;
Matches 242; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 15 KVPKIQIVOSGPELKKPEETKISCKASGYPTNYGNMVKQAPGGGLKMMGWINTSGSTF 74
DB 29 KLASGVQIQOSGPELKKPEETKISCKASGYPTNYGNMVKQAPGGGLKMMGWINTSGST 88
QY 75 ESTRADPKRFPDSLETSTANTAVLIQINNLKSEDMATYFCARMEVYHGVYPWGGTIVT 134
DB 89 ESTRADPKRFPDSLETSTANTAVLIQINNLKSEDMATYFCARMEVYHGVYPWGGTIVT 148
QY 135 VSSGGGGGGGGGGSDIQLTQSHKFLSTSVCDRYSITCKASQDVYNAVANYQKPEGO 194
DB 149 VSSGGGGGGGGGGSDIQLTQSHKFLSTSVCDRYSITCKASQDVYNAVANYQKPEGO 208
QY 195 SPKLLIYSASSRRTYGVSRFTSGSGGPDFTTISVQAEADLAYVFCQOHRTPTTSGST 254
DB 209 SPKLLIYSASSRRTYGVSRFTSGSGGPDFTTISVQAEADLAYVFCQOHRTPTTSGST 268
QY 255 KLEIKALE 262
DB 269 KLEIKALE 276

```

```

RESULT 8
AAR26980
ID AAR26980 standard; protein; 711 AA.
XX AAR26980;
AC 25-MAR-2003 (revised)
XX 11-FEB-1993 (first entry)
DE Fv(FRP5)-phoA recombinant antibody.
XX Monoclonal antibody; light chain; heavy chain; tumour; phoA;
XX variable region; alkaline phosphatase; c-erbB-2; ss.
XX Synthetic.
XX Key Location/Qualifiers
FH Peptide 1..21
FT /label= ompa_signal_peptide
FT Domain 23..141
FT /label= FRP5_heavy_chain_variable_domain
FT Peptide 142..156
FT /label= Linker
FT Domain 157..264
FT /label= FRP5_light_chain_variable_domain
FT Protein 265..711
FT /label= phoA
XX EP502812-A1.
XX 09-SEP-1992.
XX 27-JAN-1992; 92EP-00810056.
XX 05-FEB-1991; 91EP-00810079.
XX (CIBA ) CIBA GEIGY AG.
XX Wels WS, Hynes NE, Harwerth I, Groner B, Hardman N, Zwickl M;
XX WPI: 1992-302096/37.
XX N-PSDB; AAQ28257.
XX Recombinant antibodies directed to growth factor receptor C-ERBB-2 - for
XX diagnosing and treating tumours expressing C-ERBB-2 e.g. breast or
XX ovarian tumours.
XX PS Disclosure: Page 34-40; 67pp; English.
XX The sequence given is the single chain recombinant antibody designated
XX Fv(FRP5)-phoA. The alkaline phosphatase gene (phoA) was used as a marker
XX gene so that E. coli transformed with the fusion gene could be
XX identified. The fusion gene was expressed in E. coli and the antibody was
XX extracted. This recombinant antibody can be used for the qualitative and
XX quantitative determination of c-erbB-2. This can be used for monitoring
XX or in-vivo localisation of tumours overexpressing c-erbB-2. (Updated on
XX 25-MAR-2003 to correct PN field.)
XX Sequence 711 AA;

```

```

Query Match 53.4%; Score 1303; DB 2; Length 711;
Best Local Similarity 94.2%; Pred. No. 1.2e-81;
Matches 245; Conservative 2; Mismatches 1; Indels 12; Gaps 1;
QY 19 QIVQVOSGPELKKPEETKISCKASGYPTNYGNMVKQAPGGGLKMMGWINTSGSTF 78
DB 23 QIVQVOSGPELKKPEETKISCKASGYPTNYGNMVKQAPGGGLKMMGWINTSGSTF 82
QY 79 ADDPKGRFPDSLETSTANTAVLIQINNLKSEDMATYFCARMEVYHGVYPWGGTIVTSSG 138
DB 83 ADDPKGRFPDSLETSTANTAVLIQINNLKSEDMATYFCARMEVYHGVYPWGGTIVTSSG 142

```

```

QY 139 GGGSGGGGGGGSDIQLTQSHKFLSTSVGDRVSITCKASQDVYNAVAVMYOQKPGQSPKL 198
DB 143 GGGSGGGGGGGSDIQLTQSHKFLSTSVGDRVSITCKASQDVYNAVAVMYOQKPGQSPKL 202
QY 199 LIYSASSRRTGVPSRFRTSGSGSPDFTFTISSVOAEDLAVYFCQGHFRTPFTGSGTKLEI 258
DB 203 LIYSASSRRTGVPSRFRTSGSGSPDFTFTISSVOAEDLAVYFCQGHFRTPFTGSGTKLEI 262
QY 259 KALEISNSVMVFSSVAVPVLQ 278
DB 263 KALE-----PVLE 270

RESULT 9
AAR95053 standard; protein; 530 AA.
ID AAR95053
AC AAR95053;
XX 16-OCT-2003 (revised)
DT 18-AUG-1996 (first entry)
XX scFv (FRP5)-DETA-DGAL4 multidomain protein.
DE scFv (FRP5)-DETA-DGAL4 multidomain protein.
XX Nucleic acid transfer system; gene transfer; gene therapy;
KW cell targeting; multidomain protein; vector; cancer; exotoxin A; DETA;
KM single chain antibody; scFv; GAL4.
XX
OS Mus; sp.
OS Pseudomonas; aeruginosa.
OS Saccharomyces cerevisiae.
OS Chimeric.
OS Synthetic.
XX
FH Location/Qualifiers
FT 1..8
FT /label= FLAG_epitope
FT 9..17
FT /label= Spacer
FT 18..257
FT /label= ScFv (FRP5)
FT 258..260
FT /label= Spacer
FT 261..375
FT /label= ETR
FT /note= "amino acids 252-366 of exotoxin-A"
FT Peptide
FT 376
FT /label= Spacer
FT 377..522
FT /label= GAL4
FT /note= "amino acids 2-147 of yeast GAL4"
FT Peptide
FT 523..530
FT /label= Spacer
FT /note= "endoplasmic reticulum retention peptide"
XX
FN WO9613599-A1.
XX
PD 09-MAY-1996.
XX
PF 31-OCT-1995; 95WO-EP004270.
XX
PR 01-NOV-1994; 94BP-00810627.
XX
PA (WEIS/) WEIS W.
XX
PI Weis W, Fominaya J;
XX
DR MPI; 1996-239505/24.
XX
DR N-PSDB; AAT29409.
XX
PT Nucleic acid transfer system for gene therapy, e.g. against cancer -
PT includes toxin translocation domain to target nucleic acid to specific
PT cell.

```

```

XX
PS Claim 7; Page 59-61; 106pp; English.
XX
CC A multidomain protein (AAR95053) has a FLAG epitope, a single chain
CC antibody, scFv, of monoclonal antibody FRP5 (raised against human tumour
CC cell HBR2 antigen) that acts as a ligand domain, a non-cytotoxic portion
CC of Pseudomonas aeruginosa exotoxin A acting as a translocation domain and
CC the DNA binding domain of yeast GAL4. It is the product of a fusion gene
CC (AAT29409) and can be expressed in E. coli (resulting in removal of ompA
CC signal peptide). It is used with an effector nucleic acid that comprises
CC e.g. a gene to be delivered to a cell and a cognate structure for the
CC GAL4 DNA binding domain. This provides a novel means of nucleic acid
CC transfer, suitable for gene therapy. (Updated on 16-OCT-2003 to
CC standardise OS field)
XX
SQ Sequence 530 AA;
XX
Query Match 53.3%; Score 1302; DB 2; Length 530;
Best Local Similarity 99.2%; Pred. No. 1e-81;
Matches 242; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 19 QIQLVQSGPELKRPETVKISCKASGYPTVYGMWVQAQGGILKMGWINTSGESTF 78
DB 18 QVQLQDSGPDLKPKPEYVKISCKASGYPTVYGMWVQAQGGILKMGWINTSGESTF 77
QY 79 ADDPKGRDFSLERTSANTAYIQLNNLKSEDMATYFCARMEYHGVVPYMGQTTVTVSSG 138
DB 78 ADDPKGRDFSLERTSANTAYIQLNNLKSEDMATYFCARMEYHGVVPYMGQTTVTVSSG 137
QY 139 GGGSGGGGGGGSDIQLTQSHKFLSTSVGDRVSITCKASQDVYNAVAVMYOQKPGQSPKL 198
DB 138 GGGSGGGGGGGSDIQLTQSHKFLSTSVGDRVSITCKASQDVYNAVAVMYOQKPGQSPKL 197
QY 199 LIYSASSRRTGVPSRFRTSGSGSPDFTFTISSVOAEDLAVYFCQGHFRTPFTGSGTKLEI 258
DB 198 LIYSASSRRTGVPSRFRTSGSGSPDFTFTISSVOAEDLAVYFCQGHFRTPFTGSGTKLEI 257
QY 259 KALE 262
DB 258 KALE 261

RESULT 10
AAM05136 standard; protein; 651 AA.
ID AAM05136
AC AAM05136;
XX
XX 29-JAN-1997 (first entry)
DT 29-JAN-1997
XX scFv (FRP5)-ETRA fusion protein.
DE scFv (FRP5)-ETRA fusion protein.
XX
XX Single chain antibody; scFv; monoclonal antibody; Mab; erbB-2;
KW epidermal growth factor; receptor; plasmid pSW202-5; cancer; therapy;
KM antitumour; exotoxin A; ETRA.
XX
OS Mus; sp.
OS Pseudomonas; aeruginosa.
OS Synthetic.
OS Chimeric.
XX
FH Location/Qualifiers
FT 1..21
FT /label= Sig_peptide
FT /note= "ompA signal peptide"
FT 22..38
FT /label= Spacer
FT 39..278
FT /label= scFv (FRP5)
FT 279..289
FT /label= Spacer
FT 290..651
FT /label= ETRA
FT Protein

```

/note= "exotoxin A amino acids 252-613"

EP739984-A1.

30-OCT-1996.

26-APR-1995; 95EP-00106275.

26-APR-1995; 95EP-00106275.

(SANT-) SAN TUMORFORSCHUNGS GMBH.

Wels W, Schmidt M, Groner B;

WPI; 1996-478748/48.

N-PSDB; AAT42036.

Bivalent fusion proteins that bind epidermal growth factor receptor or analogues - and comprise at least two different cell surface binding domain(s), useful for tumour therapy.

Example 7; Page 21-22; 52pp; English.

scFv(FRP5)-ETA (AAW05136) comprises the single-chain binding region (see also AAW05134) of murine monoclonal antibody FRP05, which is specific for human epidermal growth factor receptor erbB-2, joined to exotoxin A (ETA). It is encoded by plasmid pSW202-5 (see also AAT42036) obtd. by ligating an scFv(FRP5) gene (AAT42034) into plasmid pSW200 contg. the Pseudomonas aeruginosa PAK ETA gene. The construct can be used to produce novel bivalent fusion proteins (see also AAW05135-44) in bacterial host cells, for use as antitumour agents

Sequence 651 AA:

Query Match 53.3%; Score 1302; DB 2; Length 651;
Best Local Similarity 99.2%; Pred. No. 1.3e-81;
Matches 242; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

19 QIQVOSGELKKPEETVYKISCKASGYPTNYGMNWKQAPGQGLKMWGINTSTGSETF 78
19 QVQLQSGELKKPEETVYKISCKASGYPTNYGMNWKQAPGQGLKMWGINTSTGSETF 98
79 ADDFKGRFDPSLETSANTAYVLIQINLKSEDMATYFCARMEYHGVHPYWGQGTIVTVSSG 138
99 ADDFKGRFDPSLETSANTAYVLIQINLKSEDMATYFCARMEYHGVHPYWGQGTIVTVSSG 158
139 GGGSGGGSGGGSDIQLTQSHKFLSTVGVDRVSTTCASQDVYNAVAMVQKFGQSEFKL 198
159 GGGSGGGSGGGSDIQLTQSHKFLSTVGVDRVSTTCASQDVYNAVAMVQKFGQSEFKL 218
199 LIYSASSRYTVPRSFRTGSGGPDFTTSSVQAEADLAIVFCQGHFRTPTFGSGTKLEI 258
219 LIYSASSRYTVPRSFRTGSGGPDFTTSSVQAEADLAIVFCQGHFRTPTFGSGTKLEI 278
259 KALE 262
279 KALE 282

RESULT 11
AAW05138
ID AAW05138 standard; protein; 699 AA.

AAW05138;
29-JAN-1997 (first entry)
scFv(FRP5)/TGF alpha-ETA.

Single chain antibody; scFv; monoclonal antibody; Mab; EGF; erbB-2;
epidermal growth factor; receptor; plasmid pMS238-5-TGF; cancer;
exotoxin A; ETA; transforming growth factor alpha; TGF; antitumour.

OS Homo, sapiens.
OS Mus; sp.
OS Pseudomonas; aeruginosa.
OS Synthetic.
OS Chimeric.

Key Location/Qualifiers
FT Peptide 1..21
FT Peptide /label= Sig peptide
FT Peptide /note= "ompa signal peptide"
FT Peptide 22..38
FT Region /label= Spacer
FT Region 39..278
FT Peptide /label= scFv(FRP5)
FT Peptide 279..289
FT Region /label= Spacer
FT Region 290..404
FT Peptide /label= ETA
FT Peptide /note= "exotoxin A amino acids 252-366"
FT Peptide 405..410
FT Region /label= Spacer
FT Region 411..460
FT Peptide /label= TGF-alpha
FT Peptide 461..465
FT Region /label= Spacer
FT Region 466..699
FT Region /label= ETA
FT Region /note= "exotoxin A amino acids 380-613"

EP739984-A1.

30-OCT-1996.

26-APR-1995; 95EP-00106275.

26-APR-1995; 95EP-00106275.

(SANT-) SAN TUMORFORSCHUNGS GMBH.

Wels W, Schmidt M, Groner B;

WPI; 1996-478748/48.

N-PSDB; AAT42038.

Bivalent fusion proteins that bind epidermal growth factor receptor or analogues - and comprise at least two different cell surface binding domain(s), useful for tumour therapy.

Example 10; Page 25-27; 52pp; English.

scFv(FRP5)/TGF alpha-ETA (AAW05138) comprises the single-chain binding region of murine monoclonal antibody FRP5 (specific for human epidermal growth factor receptor erbB-2, see also AAW05134) joined to portions of exotoxin A from Pseudomonas aeruginosa and to human transforming growth factor (TGF) alpha (see also AAW05137). It is encoded by plasmid pMS238-5-TGF (AAT42038). This plasmid can be utilized in the prodn. of the CC bispecific fusion protein in bacterial (esp. E. coli) host cells. Such fusion proteins (see also AAW05139-44) are useful as antitumour agents

Sequence 699 AA:

Query Match 53.3%; Score 1302; DB 2; Length 699;
Best Local Similarity 99.2%; Pred. No. 1.4e-81;
Matches 242; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

19 QIQVOSGELKKPEETVYKISCKASGYPTNYGMNWKQAPGQGLKMWGINTSTGSETF 78
19 QVQLQSGELKKPEETVYKISCKASGYPTNYGMNWKQAPGQGLKMWGINTSTGSETF 98
79 ADDFKGRFDPSLETSANTAYVLIQINLKSEDMATYFCARMEYHGVHPYWGQGTIVTVSSG 138
99 ADDFKGRFDPSLETSANTAYVLIQINLKSEDMATYFCARMEYHGVHPYWGQGTIVTVSSG 158

```

Oy 139 GGGSGGGSGGGSDIQLTOSHKFLSTSVGDRVSIITCKASQDVYNAVAMVYQKPGQSPKL 198
Db 159 GGGSGGGSGGGSDIQLTOSHKFLSTSVGDRVSIITCKASQDVYNAVAMVYQKPGQSPKL 218
Oy 199 LIYSASSRYTGVPSRFITGSGGPDFTFTISSVQAEDLAVYFCQGHFRPTFTGSGTKLEI 258
Db 219 LIYSASSRYTGVPSRFITGSGGPDFTFTISSVQAEDLAVYFCQGHFRPTFTGSGTKLEI 278
Oy 259 KALE 262
Db 279 KALE 282

RESULT 12
AAW05140
ID AAW05140 standard; protein; 892 AA.
AC AAW05140;
XX 29-JAN-1997 (first entry)
DE scFv2(225/FRP5)-ETA.
XX
XX Single chain antibody; scFv; monoclonal antibody; Mab; EGF; erbB-2;
KW epidermal growth factor; receptor; plasmid pMS338-225-5; cancer;
KM exotoxin A; ETA; antitumour.
XX
XX Mus; SP.
OS Pseudomonas; aeruginosa.
OS Synthetic.
OS Chimeric.
XX
FH Key
FH Peptide
FT 1..21 Location/Qualifiers
FT /label= Sig_peptide
FT /note= "ompa signal peptide"
FT Peptide
FT 22..38
FT /label= Spacer
FT Region
FT 39..278
FT /label= scFv(225)
FT Peptide
FT 279..289
FT /label= Spacer
FT Region
FT 290..404
FT /label= ETA
FT /note= "exotoxin A amino acids 252-366"
FT Peptide
FT 405..407
FT /label= Spacer
FT Region
FT 408..647
FT /label= scFv(FRP5)
FT Peptide
FT 648..658
FT /label= Spacer
FT Region
FT 659..892
FT /label= ETA
FT /note= "endotoxin-A amino acids 380-613"
XX
XX EP739984-A1.
XX
XX 30-OCT-1996.
XX
XX 26-APR-1995; 95EP-00106275.
XX
XX 26-APR-1995; 95EP-00106275.
XX
XX (SANT-) SAN TUMORFORSCHUNGS GMBH.
XX
XX Wels W, Schmidt M, Groner B;
XX
XX WPI; 1996-478748/48.
XX
XX N-P8DB; AAT42040.
XX
XX Bivalent fusion proteins that bind epidermal growth factor receptor or
PT analogues - and comprise at least two different cell surface binding
PT domain(s), useful for tumour therapy.

```

```

XX
PS Example 11; Page 31-33; 52pp; English.
XX
CC scFv2(225/FRP5)-ETA (AAW05140) comprises the single-chain binding region
CC of murine monoclonal antibody 225 (specific for human epidermal growth
CC factor receptor, see also AAW05133) joined to portions of exotoxin A from
CC Pseudomonas aeruginosa and to the single-chain binding region of murine
CC monoclonal antibody FRP5 (specific for human epidermal growth factor
CC receptor erbB-2, see also AAW05134). It is encoded by plasmid pMS238-225-
CC 5 (AAT42040). This plasmid can be utilised in the prodn. of the bivalent
CC fusion protein in bacterial (esp. E. coli) host cells. Such fusion
CC proteins (see also AAW05138-44) are useful as antitumour agents
XX
XX
SQ Sequence 892 AA;
XX
Query Match 53.3%; Score 1302; DB 2; Length 892;
Best Local Similarity 99.2%; Pred. No. 1,8e-81;
Matches 242; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Oy 19 QIQLVQSGPELKKPGETVYKISCKASGYPTNYGMNMYVQAQGLKMMGWINTSGESTF 78
Db 408 QVQLQQSGPELKKPGETVYKISCKASGYPTNYGMNMYVQAQGLKMMGWINTSGESTF 467
Oy 79 ADDEKGRFDSLETSANTAYVQINNLSKEDMATYFCARWEYVHGVPYWGQTTVTSVG 138
Db 468 ADDEKGRFDSLETSANTAYVQINNLSKEDMATYFCARWEYVHGVPYWGQTTVTSVG 527
Oy 139 GGGSGGGSGGGSDIQLTOSHKFLSTSVGDRVSIITCKASQDVYNAVAMVYQKPGQSPKL 198
Db 528 GGGSGGGSGGGSDIQLTOSHKFLSTSVGDRVSIITCKASQDVYNAVAMVYQKPGQSPKL 587
Oy 199 LIYSASSRYTGVPSRFITGSGGPDFTFTISSVQAEDLAVYFCQGHFRPTFTGSGTKLEI 258
Db 588 LIYSASSRYTGVPSRFITGSGGPDFTFTISSVQAEDLAVYFCQGHFRPTFTGSGTKLEI 647
Oy 259 KALE 262
Db 648 KALE 651

RESULT 13
AAW05143
ID AAW05143 standard; protein; 892 AA.
XX
XX AAW05143;
AC
XX
XX 29-JAN-1997 (first entry)
DE scFv2(FRP5/FRP5)-ETA (version 2).
XX
XX Single chain antibody; scFv; monoclonal antibody; Mab; EGF; erbB-2;
KW epidermal growth factor; receptor; plasmid pMS338-5-5; cancer;
KM exotoxin A; ETA; antitumour.
XX
XX Mus; SP.
OS Pseudomonas; aeruginosa.
OS Synthetic.
OS Chimeric.
XX
FH Key
FH Peptide
FT 1..21 Location/Qualifiers
FT /label= Sig_peptide
FT /note= "ompa signal peptide"
FT Peptide
FT 22..38
FT /label= Spacer
FT Region
FT 39..278
FT /label= scFv(FRP5)
FT Peptide
FT 279..289
FT /label= Spacer
FT Region
FT 290..404
FT /label= ETA
FT /note= "exotoxin A amino acids 252-366"
FT Peptide
FT 405..407

```

```

FT      /label= Spacer
FT      408..647
FT      /label= scFv(FRP5)
FT      Peptide
FT      648..658
FT      /label= Spacer
FT      Region
FT      659..892
FT      /label= ETA
FT      /note= "exotoxin A amino acids 380-613"
PN      EP739984-A1.
XX      30-OCT-1996.
XX      26-APR-1995; 95EP-00106275.
XX      26-APR-1995; 95EP-00106275.
XX      (SANT-) SAN TUMORFORSCHUNGS GMBH.
XX      Wels W, Schmidt M, Groner B;
XX      MPI; 1996-478748/48.
XX      N-PSDB; AAT42043.
XX      Bivalent fusion proteins that bind epidermal growth factor receptor or
XX      PT analogues - and comprise at least two different cell surface binding
XX      PT domain(s), useful for tumour therapy.
XX      Example 12; Page 40-42; 52pp; English.
XX      scFv2(FRP5/FRP5)-ETA version 2 (AAW05139) includes 2 copies of the single
XX      chain binding region of murine monoclonal antibody FRP5 (specific for
XX      human epidermal growth factor receptor erbB-2, see also AAW05134) joined
XX      to portions of exotoxin A from Pseudomonas aeruginosa. It is encoded by
XX      CC plasmid pMS238-5-5 (AAT42043). This plasmid can be utilised in the prodn.
XX      CC of the bivalent fusion protein in bacterial (esp. E. coli) host cells.
XX      CC Such fusion proteins (see also AAW05138-44) are useful as antitumour
XX      agents
XX      SQ      Sequence 892 AA;
XX
Query Match      53.3%; Score 1302; DB 2; Length 892;
Best Local Similarity 99.2%; Pred. No. 1.8e-81;
Matches 242; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY      19 QIQLVQSGPELKKPEETVYKISCKASGYPTNYGMNWKQAPGQGLKMMGWINTSTGESTF 78
DB      39 QVQLQOQSGPELKKPEETVYKISCKASGYPTNYGMNWKQAPGQGLKMMGWINTSTGESTF 98
QY      79 ADDFKGRFDESLSTANTAYLQINNLSKEDMATYTCANWEYHGVVPPYWGQGTIVTVSSG 138
DB      99 ADDFKGRFDESLSTANTAYLQINNLSKEDMATYTCANWEYHGVVPPYWGQGTIVTVSSG 158
QY      139 GGGSGGGSGGGGSDIQLTQSHKFLSTSVGDRVSTTCASQDVYNAVAWYQKFGQSPKL 198
DB      159 GGGSGGGSGGGGSDIQLTQSHKFLSTSVGDRVSTTCASQDVYNAVAWYQKFGQSPKL 218
QY      199 LIYSASSYVTGWPSPFTSSGSGPDTFTISSVQADLAIVFCQGHFPTFTFGSGTLEI 258
DB      219 LIYSASSYVTGWPSPFTSSGSGPDTFTISSVQADLAIVFCQGHFPTFTFGSGTLEI 278
QY      259 KALE 262
DB      279 KALE 282

```

RESULT 14
AAW05139
ID AAW05139 standard; protein; 892 AA.

XX AC AAW05139;
XX DT 29-JAN-1997 (first entry)

```

XX      scFv2(FRP5/225)-ETA (version 1).
DE      Single chain antibody; scFv; monoclonal antibody; Mab; EGF; erbB-2;
XX      epidermal growth factor; receptor; plasmid pMS238-5-225; cancer;
XX      exotoxin A; ETA; antitumour.
XX      Mus; sp.
XX      Pseudomonas; aeruginosa.
XX      Synthetic.
XX      Chimeric.
XX      Key
XX      Peptide
XX      Location/Qualifiers
XX      1..21
XX      /label= Sig peptide
XX      /note= "ompA signal peptide"
XX      22..38
XX      /label= Spacer
XX      39..278
XX      /label= scFv(FRP5)
XX      279..289
XX      /label= Spacer
XX      290..404
XX      /label= ETA
XX      /note= "exotoxin A amino acids 252-366"
XX      405..407
XX      /label= Spacer
XX      408..647
XX      /label= scFv(225)
XX      648..658
XX      /label= Spacer
XX      659..892
XX      /label= ETA
XX      /note= "endotoxin-A amino acids 380-613"
XX      EP739984-A1.
XX      30-OCT-1996.
XX      26-APR-1995; 95EP-00106275.
XX      26-APR-1995; 95EP-00106275.
XX      (SANT-) SAN TUMORFORSCHUNGS GMBH.
XX      Wels W, Schmidt M, Groner B;
XX      MPI; 1996-478748/48.
XX      N-PSDB; AAT42039.
XX      Bivalent fusion proteins that bind epidermal growth factor receptor or
XX      PT analogues - and comprise at least two different cell surface binding
XX      PT domain(s), useful for tumour therapy.
XX      Example 11; Page 28-30; 52pp; English.
XX      scFv2(FRP5/225)-ETA (AAW05139) comprises the single-chain binding region
XX      of murine monoclonal antibody FRP5 (specific for human epidermal growth
XX      factor receptor erbB-2, see also AAW05134) joined to portions of exotoxin
XX      A from Pseudomonas aeruginosa and to the single-chain binding region of
XX      murine monoclonal antibody 225 (specific for human epidermal growth
XX      factor receptor, see also AAW05133). It is encoded by plasmid pMS238-5-
XX      225 (AAT42039). This plasmid can be utilised in the prodn. of the
XX      CC bivalent fusion protein in bacterial (esp. E. coli) host cells. Such
XX      CC fusion proteins (see also AAW05138-44) are useful as antitumour agents
XX      SQ      Sequence 892 AA;
XX
Query Match      53.3%; Score 1302; DB 2; Length 892;
Best Local Similarity 99.2%; Pred. No. 1.8e-81;
Matches 242; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY      19 QIQLVQSGPELKKPEETVYKISCKASGYPTNYGMNWKQAPGQGLKMMGWINTSTGESTF 78

```

```

DB 39 QVQLQSGPELKKPGETYKISCKASGYPTNYGMNWMVQAQPGGLKMMGMINTSTGSESTF 98
QY 79 ADDPKGRDPSLETSANTAYLQINNLSKEDMATYFCARMEVYHGVPYWGQGTITVVS 138
DB 99 ADDPKGRDPSLETSANTAYLQINNLSKEDMATYFCARMEVYHGVPYWGQGTITVVS 158
QY 139 GGGSGGGSGGGSGSDIQLTQSHKFLSTSVGDRVSIITCKASQDVYNAVAMVYQKPGQSPKL 198
DB 159 GGGSGGGSGGGSGSDIQLTQSHKFLSTSVGDRVSIITCKASQDVYNAVAMVYQKPGQSPKL 218
QY 199 LIYSASSRYTGVPSRFTSGSGSDPFTTISVQAEADLAVYFCQGHFRTPFTFGSGTKLEI 258
DB 219 LIYSASSRYTGVPSRFTSGSGSDPFTTISVQAEADLAVYFCQGHFRTPFTFGSGTKLEI 278
QY 259 KALE 262
DB 279 KALE 282

```

RESULT 15

```

AAW05142 ID AAW05142 standard; protein; 895 AA.
XX AC AAW05142;
XX DT 29-JAN-1997 (first entry)
XX DE scFv2(FRP5)/ETA (version 1).
XX KW Single chain antibody; scFv; monoclonal antibody; Mab; EGF; erbB-2;
KM epidermal growth factor; receptor; plasmid pMS242-5-5; cancer;
XX exotoxin A; ETA; antitumour.
XX MS; sp.
OS Pseudomonas; aeruginosa.
OS Synthetic.
OS Chimeric.
XX FH Key
FH FT Location/Qualifiers
FH FT 1..21
FH FT /label= Sig_peptide
FH FT /note= "ompA signal peptide"
FH FT Peptide
FH FT 22..44
FH FT /label= Spacer
FH FT Region
FH FT 45..159
FH FT /label= ETA
FH FT /note= "exotoxin A amino acids 252-366"
FH FT Peptide
FH FT 160..162
FH FT /label= Spacer
FH FT Region
FH FT 163..402
FH FT /label= scFv(FRP5)
FH FT Peptide
FH FT 403..413
FH FT /label= Spacer
FH FT Region
FH FT 414..644
FH FT /label= ETA
FH FT /note= "exotoxin A amino acids 380-610"
FH FT Peptide
FH FT 645..646
FH FT /label= Spacer
FH FT Region
FH FT 647..886
FH FT /label= scFv(FRP5)
FH FT Peptide
FH FT 887..895
FH FT /label= Spacer
XX PN EP73984-A1.
XX PD 30-OCT-1996.
XX PF 26-APR-1995; 95EP-00106275.
XX PR 26-APR-1995; 95EP-00106275.
XX PA (SANT-) SAN TUMORFORSCHUNGS GMBH.

```

```

XX PI Weis W, Schmidt M, Groner B;
XX DR WPI: 1996-478746/48.
XX DR N-PSDB: AAT42042.
XX PT Bivalent fusion proteins that bind epidermal growth factor receptor or
XX PT analogues - and comprise at least two different cell surface binding
XX PT domain(s), useful for tumour therapy.
XX PS Example 12: Page 37-39; 52pp; English.
XX CC scFv2(FRP5)/ETA version 1 (AAW05142) includes 2 copies of the single
XX CC -chain binding region of murine monoclonal antibody FRP5 (specific for
XX CC human epidermal growth factor receptor erbB-2, see also AAW05134) joined
XX CC to portions of exotoxin A from Pseudomonas aeruginosa. It is encoded by
XX CC plasmid pMS242-5-5 (AAT42042). This plasmid can be utilised in the prodn.
XX CC of the bivalent fusion protein in bacterial (esp. E. coli) host cells.
XX CC Such fusion proteins (see also AAW05138-44) are useful as antitumour
XX CC agents
XX SQ Sequence 895 AA:

```

```

Query Match 53.3%; Score 1302; DB 2; Length 895;
Best Local Similarity 99.2%; Pred. No. 1.8e-81;
Matches 242; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 19 QIQLVQSGPELKKPGETYKISCKASGYPTNYGMNWMVQAQPGGLKMMGMINTSTGSESTF 78
DB 163 QVQLQSGPELKKPGETYKISCKASGYPTNYGMNWMVQAQPGGLKMMGMINTSTGSESTF 222
QY 79 ADDPKGRDPSLETSANTAYLQINNLSKEDMATYFCARMEVYHGVPYWGQGTITVVS 138
DB 223 ADDPKGRDPSLETSANTAYLQINNLSKEDMATYFCARMEVYHGVPYWGQGTITVVS 282
QY 139 GGGSGGGSGGGSGSDIQLTQSHKFLSTSVGDRVSIITCKASQDVYNAVAMVYQKPGQSPKL 198
DB 283 GGGSGGGSGGGSGSDIQLTQSHKFLSTSVGDRVSIITCKASQDVYNAVAMVYQKPGQSPKL 342
QY 199 LIYSASSRYTGVPSRFTSGSGSDPFTTISVQAEADLAVYFCQGHFRTPFTFGSGTKLEI 258
DB 343 LIYSASSRYTGVPSRFTSGSGSDPFTTISVQAEADLAVYFCQGHFRTPFTFGSGTKLEI 402
QY 259 KALE 262
DB 403 KALE 406

```

```

Search completed: April 20, 2005, 07:04:52
Job time : 70.8241 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2005, 07:01:08 ; Search time 20.9082 Seconds
(without alignments)
1645.921 Million cell updates/sec

Title: US-09-596-774-6
Perfect score: 2441
Sequence: 1 MAWWTLLFLMAAKVPKQI.....ISTATKDTYDALHMQTLLAPR 461

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues
Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	1308	53.6	615 4 US-08-840-713-35	Sequence 35, Appl
2	1308	53.6	617 4 US-08-840-713-37	Sequence 37, Appl
3	1307	53.5	637 1 US-08-235-838-14	Sequence 14, Appl
4	1307	53.5	637 2 US-08-465-473B-14	Sequence 14, Appl
5	1303	53.4	711 1 US-08-235-838-7	Sequence 7, Appl
6	1303	53.4	711 2 US-08-465-473B-7	Sequence 7, Appl
7	1302	53.3	530 4 US-08-840-713-2	Sequence 2, Appl
8	1284	52.6	241 1 US-08-235-838-5	Sequence 5, Appl
9	1284	52.6	241 2 US-08-465-473B-5	Sequence 5, Appl
10	981.5	40.2	250 1 US-08-133-804-2	Sequence 8, Appl
11	981.5	40.2	250 1 US-08-461-184-8	Sequence 8, Appl
12	981.5	40.2	250 1 US-08-463-675-8	Sequence 8, Appl
13	981.5	40.2	250 1 US-08-464-589-8	Sequence 8, Appl
14	981.5	40.2	250 1 US-08-461-838-2	Sequence 2, Appl
15	981.5	40.2	250 2 US-08-461-386-2	Sequence 2, Appl
16	975.5	40.0	622 2 US-08-356-786-16	Sequence 16, Appl
17	903.5	37.0	240 1 US-08-488-113B-148	Sequence 148, App
18	903.5	37.0	240 1 US-08-477-488B-148	Sequence 148, App
19	903.5	37.0	240 2 US-08-466-360-148	Sequence 148, App
20	903.5	37.0	240 3 US-08-839-765-148	Sequence 148, App
21	903.5	37.0	240 3 US-09-136-389-148	Sequence 148, App
22	903.5	37.0	240 3 US-09-610-838-148	Sequence 148, App
23	903.5	37.0	240 4 US-09-711-485-148	Sequence 148, App
24	901.5	36.9	267 3 US-09-485-737B-2	Sequence 2, Appl
25	901.5	36.9	267 4 US-10-071-485-2	Sequence 2, Appl
26	896	36.7	365 3 US-08-875-811-53	Sequence 53, Appl
27	891	36.5	366 3 US-08-875-811-55	Sequence 55, Appl

28	889.5	36.4	541 3 US-09-485-737B-85	Sequence 85, Appl
29	889.5	36.4	541 4 US-10-071-485-85	Sequence 85, Appl
30	888.5	36.4	711 3 US-09-485-737B-90	Sequence 90, Appl
31	888.5	36.4	711 4 US-10-071-485-90	Sequence 90, Appl
32	888	36.4	284 3 US-09-184-658-40	Sequence 40, Appl
33	888	36.4	284 4 US-09-504-262D-40	Sequence 40, Appl
34	885.5	36.3	259 4 US-09-419-788-115	Sequence 115, App
35	878.5	36.0	534 2 US-08-356-786-10	Sequence 10, Appl
36	851	34.9	240 4 US-10-092-246-35	Sequence 35, Appl
37	851	34.9	240 4 US-10-092-246-36	Sequence 36, Appl
38	851	34.9	240 4 US-10-096-246A-35	Sequence 35, Appl
39	851	34.9	240 4 US-10-096-246A-37	Sequence 37, Appl
40	848	34.7	240 3 US-09-485-737B-91	Sequence 91, Appl
41	848	34.7	240 4 US-10-071-485-91	Sequence 91, Appl
42	843	34.5	240 4 US-10-096-246A-36	Sequence 36, Appl
43	841.5	34.5	249 2 US-08-797-689-18	Sequence 18, Appl
44	841.5	34.5	249 4 US-09-984-186-18	Sequence 18, Appl
45	839	34.4	240 4 US-10-092-246-37	Sequence 37, Appl

ALIGNMENTS

```

RESULT 1
US-08-840-713-35
; Sequence 35, Application US/08840713
; Patent No. 6498233
;
GENERAL INFORMATION:
; APPLICANT: WELS, Winfried, Dr.
; APPLICANT: FOYMINAYA, Jesus
; TITLE OF INVENTION: NUCLEIC ACID TRANSFER SYSTEM
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESS: Nikaido, Marmelstein, Murray & Oram LLP
; STREET: 655 15th St., N.W., Suite 330 - G St. Lobby
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-5701
;
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/840, 713
; FILING DATE: 25-APR-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Kites, Monica Chin
; REGISTRATION NUMBER: 36,105
; REFERENCE/DOCKET NUMBER: 1614-7014
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 638 - 5000
; TELEFAX: (202) 638 - 4810
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 615 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-840-713-35
Query Match 53.6%; Score 1308; DB 4; Length 615;
Best Local Similarity 96.5%; Pred. No. 2.5e-100;
Matches 245; Conservative 3; Mismatches 6; Gaps 0;
QY 19 QIOLVSGPELKKKGGTYKISCKASGYPTNYGNWYKQAPGGLKMMGWINTSGESTF 78
DB 362 QVQLQSGPELKKKGGTYKISCKASGYPTNYGNWYKQAPGGLKMMGWINTSGESTF 421
QY 79 ADKFGKGFDFSLSTANTAYLQINNLKSEDMATYFCARWEVYHGVVPYGGGTTVTYSSG 138

```

Db 422 ADDPKGRFDSLETSANTAYLIQINNLKSEDMATYFCARWEYHGVYPYWGQTTVTSSG 481
QY 139 GGGGGGGGGGGSDIQLTQSHKFLSTSVGDRVSIITCKASODVYNAVAYOQKPGQSPKL 198
Db 482 GGGGGGGGGGGSDIQLTQSHKFLSTSVGDRVSIITCKASODVYNAVAYOQKPGQSPKL 541
QY 199 LIYSASRYTGVPSRFTGSGGPDFTTISVQAEIDLAVYFCQHFRTPTFTGSGTKLEI 258
Db 542 LIYSASRYTGVPSRFTGSGGPDFTTISVQAEIDLAVYFCQHFRTPTFTGSGTKLEI 601
QY 259 KALEINSVWYFSS 272
Db 602 KALEIDLSERRRFA 615

RESULT 2

US-08-840-713-37
Sequence 37, Application US/08840713
Patent No. 6498233
GENERAL INFORMATION:
APPLICANT: WELLS, Winfried, Dr.
TITLE OF INVENTION: NOCLETIC ACID TRANSFER SYSTEM
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nikolaïdo, Matmelstein, Murray & Oram LLP
STREET: 655 15th St., N.W., Suite 330 - G St. Lobby
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-5701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/840,713
FILING DATE: 25-APR-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Kilets, Monica Chih
REGISTRATION NUMBER: 36,105
REFERENCE/DOCKET NUMBER: 1614-7014
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 638 - 4810
TELEFAX: (202) 638 - 4810
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 617 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-840-713-37

Query Match 53.6%; Score 1308; DB 4; Length 617;
Best Local Similarity 96.5%; Pred. No. 2.5e-100;
Matches 245; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 19 QIOLVQSGPELKKPGEIVYKISCKASGYPTNYGNMVKQAPGGQLKMMGMINSTSESTF 78
Db 364 QVLOQSGPELKKPGEIVYKISCKASGYPTNYGNMVKQAPGGQLKMMGMINSTSESTF 423
QY 79 ADDPKGRFDSLETSANTAYLIQINNLKSEDMATYFCARWEYHGVYPYWGQTTVTSSG 138
Db 424 ADDPKGRFDSLETSANTAYLIQINNLKSEDMATYFCARWEYHGVYPYWGQTTVTSSG 483
QY 139 GGGGGGGGGGGSDIQLTQSHKFLSTSVGDRVSIITCKASODVYNAVAYOQKPGQSPKL 198
Db 484 GGGGGGGGGGGSDIQLTQSHKFLSTSVGDRVSIITCKASODVYNAVAYOQKPGQSPKL 543
QY 199 LIYSASRYTGVPSRFTGSGGPDFTTISVQAEIDLAVYFCQHFRTPTFTGSGTKLEI 258

Db 544 LIYSASRYTGVPSRFTGSGGPDFTTISVQAEIDLAVYFCQHFRTPTFTGSGTKLEI 603
QY 259 KALEINSVWYFSS 272
Db 604 KALEIDLSERRRFA 617

RESULT 3

US-08-235-838-14
Sequence 14, Application US/08235838
Patent No. 571894
GENERAL INFORMATION:
APPLICANT: WELLS, Winfried S.
APPLICANT: Hynes, Nancy E.
APPLICANT: Harwerth, Ina-Maria
APPLICANT: Groner, Bernd
APPLICANT: Hardman, No. 5571894man
APPLICANT: Zwickl, Markus
TITLE OF INVENTION: Recombinant Antibodies Specific for a
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: New York
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235,838
FILING DATE: TBA
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/828,832
FILING DATE: 31-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 91-810079.3
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Rimer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: 4-18518/A/CIP/CONT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8614
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 637 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-235-838-14

Query Match 53.5%; Score 1307; DB 1; Length 637;
Best Local Similarity 98.0%; Pred. No. 3.1e-100;
Matches 243; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 15 KVRKQIOLVQSGPELKKPGEIVYKISCKASGYPTNYGNMVKQAPGGQLKMMGMINSTG 74
Db 29 KLSQVLOQSGPELKKPGEIVYKISCKASGYPTNYGNMVKQAPGGQLKMMGMINSTG 88
QY 75 ESTFADDFGRPFSELETSANTAYLIQINNLKSEDMATYFCARWEYHGVYPYWGQTTVT 134
Db 89 ESTFADDFGRPFSELETSANTAYLIQINNLKSEDMATYFCARWEYHGVYPYWGQTTVT 148
QY 135 VSSGGGGGGGGGGSDIQLTQSHKFLSTSVGDRVSIITCKASODVYNAVAYOQKPGQ 194
Db 149 VSSGGGGGGGGGGSDIQLTQSHKFLSTSVGDRVSIITCKASODVYNAVAYOQKPGQ 208

QY 195 SPKLIYASASSRYTGVPSRFTGSGGPDFTTISVQAEADLAIVFCQOHFRTPTFGSGT 254
 DB 209 SPKLIYASASSRYTGVPSRFTGSGGPDFTTISVQAEADLAIVFCQOHFRTPTFGSGT 268
 QY 255 KLEIKALE 262
 DB 269 KLEIKALE 276

RESULT 4
 US-08-465-473B-14
 ; Sequence 14, Application US/08465473B
 ; Patent No. 5939531

GENERAL INFORMATION:
 APPLICANT: Wels, Winfried S.
 APPLICANT: Hynes, Nancy E.
 APPLICANT: Harwerth, Ina-Maria
 APPLICANT: Groner, Bernd
 APPLICANT: Hardman, No. 5939531man
 APPLICANT: Zwickl, Markus
 TITLE OF INVENTION: Recombinant Antibodies Specific for a
 TITLE OF INVENTION: Growth Factor Receptor
 NUMBER OF SEQUENCES: 34
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: NOVARTIS Corporation
 STREET: 564 Morris Avenue
 CITY: Summit
 STATE: New Jersey
 COUNTRY: USA
 ZIP: 07901-6940
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/465,473B
 FILING DATE: 5 June 1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/828,832
 FILING DATE: 31-JAN-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 91-810079.3
 FILING DATE: 05-FEB-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Pfeiffer, Henna J.
 REGISTRATION NUMBER: 22,640
 REFERENCE/DOCKET NUMBER: 4-18518/A/CIP/CONT2
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (908)522 6940
 TELEFAX: (908)522 6955
 INFORMATION FOR SEQ ID NO: 14:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 637 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-465-473B-14

Query Match 53.5%; Score 1307; DB 2; Length 637;
 Best Local Similarity 98.0%; Pred. No. 3, 1e-100;
 Matches 243; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 15 KVPKQIQVQSGPELKKRGETVKISCKASGYPTNYGNMWWYQAPGQGLKMMGWINTSTG 74
 DB 29 KLASVQVQSGPELKKRGETVKISCKASGYPTNYGNMWWYQAPGQGLKMMGWINTSTG 88
 QY 75 ESTPADDFKGRDFSLSTANTAYVQINNLSKSEDMATYFCARWEVYHGVPYWGQTTVT 134
 DB 89 ESTPADDFKGRDFSLSTANTAYVQINNLSKSEDMATYFCARWEVYHGVPYWGQTTVT 148

QY 135 VSSGGGSGGGGSGGGSDIQLTQSHKFLSTSVGDRVSIITCKASQDVNAYVQOQKPGQ 194
 DB 149 VSSGGGSGGGGSGGGSDIQLTQSHKFLSTSVGDRVSIITCKASQDVNAYVQOQKPGQ 208
 QY 195 SPKLIYASASSRYTGVPSRFTGSGGPDFTTISVQAEADLAIVFCQOHFRTPTFGSGT 254
 DB 209 SPKLIYASASSRYTGVPSRFTGSGGPDFTTISVQAEADLAIVFCQOHFRTPTFGSGT 268
 QY 255 KLEIKALE 262
 DB 269 KLEIKALE 276

RESULT 5
 US-08-235-838-7
 ; Sequence 7, Application US/08235838
 ; Patent No. 5571894

GENERAL INFORMATION:
 APPLICANT: Wels, Winfried S.
 APPLICANT: Hynes, Nancy E.
 APPLICANT: Harwerth, Ina-Maria
 APPLICANT: Groner, Bernd
 APPLICANT: Hardman, No. 5571894man
 APPLICANT: Zwickl, Markus
 TITLE OF INVENTION: Recombinant Antibodies Specific for a
 TITLE OF INVENTION: Growth Factor Receptor
 NUMBER OF SEQUENCES: 16
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: CIBA-GEIGY Corporation
 STREET: 7 Skyline Drive
 CITY: Hawthorne
 STATE: New York
 COUNTRY: USA
 ZIP: 10532
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/235,838
 FILING DATE: TBA
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/828,832
 FILING DATE: 31-JAN-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 91-810079.3
 FILING DATE: 05-FEB-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Elmer, James Scott
 REGISTRATION NUMBER: 36,129
 REFERENCE/DOCKET NUMBER: 4-18518/A/CIP/CONT
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (919)541-8614
 TELEFAX: (919)541-8689
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 711 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-235-838-7

Query Match 53.4%; Score 1303; DB 1; Length 711;
 Best Local Similarity 94.2%; Pred. No. 7, 8e-100;
 Matches 245; Conservative 2; Mismatches 1; Indels 12; Gaps 1;

QY 19 QIQVQSGPELKKRGETVKISCKASGYPTNYGNMWWYQAPGQGLKMMGWINTSTG 78
 DB 23 QIQVQSGPELKKRGETVKISCKASGYPTNYGNMWWYQAPGQGLKMMGWINTSTG 82
 QY 79 ADPDFKGRDFSLSTANTAYVQINNLSKSEDMATYFCARWEVYHGVPYWGQTTVTYSSG 138

Db	83	ADGFKRFPDPSLETSTANTAYLQINMLKSDMATYFCARMEYHGVLPWGGTGVTVYSSG	142
Qy	139	GGGGGGGGGGSSDIQLTQSHKFLSTSVGDVSLTCAASODVYNAAMVQKFGSPKL	199
Db	143	GGGGGGGGGGSSDIQLTQSHKFLSTSVGDVSLTCAASODVYNAAMVQKFGSPKL	202
Qy	199	LIYSASRRYGVPSRREFTSGSGPDEFTTISVQAEIDLAVYFCQOHFRTPFGSGTLEI	258
Db	203	LIYSASRRYGVPSRREFTSGSGPDEFTTISVQAEIDLAVYFCQOHFRTPFGSGTLEI	262
Qy	259	KALEISNSMYFSSVVPVLQ	278
Db	263	KALE-----PVLE	270

```

US-08-465-473B-7
; Sequence 7, Application US/08465473B
; Patent No. 5939531
GENERAL INFORMATION:
APPLICANT: Wels, Winfried S.
APPLICANT: Hynes, Nancy E.
APPLICANT: Harwerth, Ina-Maria
APPLICANT: Groner, Bernd
APPLICANT: Hardman, No. 5939531man
APPLICANT: Zwickl, Markus
TITLE OF INVENTION: Recombinant Antibodies Specific for a
TITLE OF INVENTION: Growth Factor Receptor
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRES:
ADDRESSEE: NOVARTIS Corporation
STREET: 564 Morris Avenue
CITY: Summit
STATE: New Jersey
COUNTRY: USA
ZIP: 07901-6940
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,473B
FILING DATE: 5 June 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/828,832
FILING DATE: 31-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 91-810079.3
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Pfeiffer, Heena J.
REGISTRATION NUMBER: 22,640
REFERENCE/DOCKET NUMBER: 4-18518/A/CIP/CONT2
TELEPHONE: (908) 522 6940
TELEPHONE: (908) 522 6955
TELEFAX: (908) 522 6955
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 711 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-465-473B-7
Query Match 53.4%; Score 1303; DB 2; Length 711;
Best Local Similarity 94.2%; Pred. No. 7, 8e-100;
Matches 245; Conservative 2; Mismatches 1; Indels 12; Gaps
19 QIQLVSGPELKKRGRTVITISCAASGYPTNTYGNMVKQAPGCGIKMKGMINSTGSETF 789
1:|||||

```

Db	23	QVLOQSGPBLKKPBGSTVAKISCKASGSPPTNYGMMVWYKQAPQGLKMMGINTSGSIF	82
OY	79	ADDPKGRFDESLTETSANTAYLIQINNLKSEDMATYFCARMEYHGVYPMWGGITVTVSSG	138
Db	83	ADDPKGRFPFSLTETSANTAYLIQINNLKSEDMATYFCARMEYHGVYPMWGGITVTVSSG	142
OY	139	GGSGGGGGGGGGSDILOTHSHKFLSTSVGDVSTTCASQDVYNAVWYQKRGOSPKL	196
Db	143	GGSGGGGGGGGGSDILOTHSHKFLSTSVGDVSTTCASQDVYNAVWYQKRGOSPKL	202
OY	199	LIVSASRRYTGVPBRSFTSGSGSPDFTFTISSVOAEDLAVYFCQOHFRTPFTFGSGTLEI	258
Db	203	LIVSASRRYTGVPBRSFTSGSGSPDFTFTISSVOAEDLAVYFCQOHFRTPFTFGSGTLEI	262
OY	259	KALEINSNVYFSSVYVPLQ	278
Db	263	KALE-----PVLE	270

```

RESULT 7
US-08-840-713-2
; Sequence 2, Application US/08840713
; Patent No. 6498233
; GENERAL INFORMATION:
; APPLICANT: WELLS, Winfried, Dr.
; APPLICANT: FOYMINAYA, Jesus
; TITLE OF INVENTION: NUCLEIC ACID TRANSFER SYSTEM
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP
; STREET: 655 15th St., N.W., Suite 330 - G St. Lobby
; City: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MC-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/840, 713
; FILING DATE: 25-APR-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Kites, Monica Chin
; REGISTRATION NUMBER: 36,105
; REFERENCE/DOCKET NUMBER: 1614-7014
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 638 - 5000
; TELEFAX: (202) 638 - 4810
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-840-713-2

Query Match      53.3%; Score 1302; DB 4; Length 530;
Best Local Similarity 99.2%; Pred. No. 6.3e-100;
Matches 242; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 19 QIOLVSSPELKKKGGETVKISCKASGYPTNYGMNVKQAPOGGLKMGWINTSTGESTF 78
Db 18 QVOQLQQSPPELKRRGEFTVKISCKASGPFTNYGMNVKQAPGGGLKMGWINTSTGESTF 77
QY 79 ADDEKGFDFEDELFSANTAYLIQINILKSEDMATYFCARMEVYHGVPYVGSGTTVYSSG 138
Db 78 ADDRKRDPSELERSANTAYLIQINILKSEDMATYFCARMEVYHGVIPYVGSGTTVYSSG 137
QY 139 GGSGGGGGGGGGSDIQLTQSHKFLSTSVGDRAVSITCKASODVYNNAVAVYQQRPGOSPRL 198

```

Db 138 GGGSGGGGGGGSDIQLTQSHKFLSTSVGBRVISITCKASODVYNANAWYQOKPGQSPKL 197
Qy 199 LIYSASSRYTGVPSRFTGSGGSDPFTTISVQAEADLAVYFCQGHFRTPTFGSGTKLEI 258
Db 198 LIYSASSRYTGVPSRFTGSGGSDPFTTISVQAEADLAVYFCQGHFRTPTFGSGTKLEI 257
Qy 259 KALE 262
Db 258 KALE 261

RESULT 8

US-08-235-838-5
Sequence 5, Application US/08235838
Patent No. 5571894
GENERAL INFORMATION:
APPLICANT: Wels, Winfried S.
APPLICANT: Hynes, Nancy E.
APPLICANT: Harwerth, Ina-Maria
APPLICANT: Groner, Bernd
APPLICANT: Hardman, No. 5571894man
APPLICANT: Zwickl, Markus
TITLE OF INVENTION: Recombinant Antibodies Specific for a
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: New York
COUNTRY: USA
ZIP: 10512
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235,838
FILING DATE: TBA
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/828,832
FILING DATE: 31-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 91-810079.3
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Rimer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: 4-18518/A/CIP/CONT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8614
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 241 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-235-838-5

Query Match 52.6%; Score 1284; DB 1; Length 241;
Best Local Similarity 99.2%; Pred. No. 6.9e-99;
Matches 238; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 19 QIOLVOSGPELKKRPETVYKISCKASGYPTNYGNMWWQAAPGQGLKMMGMINSTSGESTF 78
Db 2 QVQLQDSGPELKKRPETVYKISCKASGYPTNYGNMWWQAAPGQGLKMMGMINSTSGESTF 61
Qy 79 ADDPKGRDPSLETSANTAYLQINNLSKEDMATYFCARWEVYHGVPYWGQTTVTYSSG 138
Db 62 ADDPKGRDPSLETSANTAYLQINNLSKEDMATYFCARWEVYHGVPYWGQTTVTYSSG 121

Qy 139 GGGSGGGGGGGSDIQLTQSHKFLSTSVGBRVISITCKASODVYNANAWYQOKPGQSPKL 198
Db 122 GGGSGGGGGGGSDIQLTQSHKFLSTSVGBRVISITCKASODVYNANAWYQOKPGQSPKL 181
Qy 199 LIYSASSRYTGVPSRFTGSGGSDPFTTISVQAEADLAVYFCQGHFRTPTFGSGTKLEI 258
Db 182 LIYSASSRYTGVPSRFTGSGGSDPFTTISVQAEADLAVYFCQGHFRTPTFGSGTKLEI 241

RESULT 9

US-08-465-473B-5
Sequence 5, Application US/08465473B
Patent No. 5939531
GENERAL INFORMATION:
APPLICANT: Wels, Winfried S.
APPLICANT: Hynes, Nancy E.
APPLICANT: Harwerth, Ina-Maria
APPLICANT: Groner, Bernd
APPLICANT: Hardman, No. 5939531man
APPLICANT: Zwickl, Markus
TITLE OF INVENTION: Recombinant Antibodies Specific for a
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: NOVARTIS Corporation
STREET: 564 Morris Avenue
CITY: Summit
STATE: New Jersey
COUNTRY: USA
ZIP: 07901-6940
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,473B
FILING DATE: 5 June 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/828,832
FILING DATE: 31-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 91-810079.3
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Pfeiffer, Henna J.
REGISTRATION NUMBER: 22,640
REFERENCE/DOCKET NUMBER: 4-18518/A/CIP/CONT2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908)522 6940
TELEFAX: (908)522 6955
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 241 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-465-473B-5

Query Match 52.6%; Score 1284; DB 2; Length 241;
Best Local Similarity 99.2%; Pred. No. 6.9e-99;
Matches 238; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 19 QIOLVOSGPELKKRPETVYKISCKASGYPTNYGNMWWQAAPGQGLKMMGMINSTSGESTF 78
Db 2 QVQLQDSGPELKKRPETVYKISCKASGYPTNYGNMWWQAAPGQGLKMMGMINSTSGESTF 61
Qy 79 ADDPKGRDPSLETSANTAYLQINNLSKEDMATYFCARWEVYHGVPYWGQTTVTYSSG 138
Db 62 ADDPKGRDPSLETSANTAYLQINNLSKEDMATYFCARWEVYHGVPYWGQTTVTYSSG 121

QY 139 GGGSGGGGGGGSDIQLTQSHKFLSTSVGDRVSITCKASQDVNAVAVYQOKPGQSPKL 198
DB 122 GGGSGGGGGGGSDIQLTQSHKFLSTSVGDRVSITCKASQDVNAVAVYQOKPGQSPKL 181
QY 199 LIYSASSRYTGVPSRFTGSGGSPDFTFTISSVQAEIDLAVYFCQGHFRPTFTFGSGTKLEI 258
DB 182 LIYSASSRYTGVPSRFTGSGGSPDFTFTISSVQAEIDLAVYFCQGHFRPTFTFGSGTKLEI 241

RESULT 10

US-08-133-804-2
Sequence 2, Application US/08133804
Patent No. 5534254
GENERAL INFORMATION:
APPLICANT: Huston, James S.
APPLICANT: Oppermann, Hermann
APPLICANT: Houston, L. L.
APPLICANT: King, David B.
TITLE OF INVENTION: Biosynthetic Binding Proteins For
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Testa, Hurwitz & Thibault/Patent Department
STREET: Exchange Place, 53 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/133,804
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Kelley, Robin D.
REGISTRATION NUMBER: 34,637
REFERENCE/DOCKET NUMBER: 2054/22
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-248-7477
TELEFAX: 617-248-7100
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 250 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-133-804-2

Query Match 40.2%; Score 981.5; DB 1; Length 250;
Best Local Similarity 76.8%; Pred. No. 1e-73;
Matches 185; Conservative 22; Mismatches 33; Indels 1; Gaps 1;

QY 19 QIQLVOSGPELKKRGFTVYKISCKASGYPTNYGMNVKQAPGQGLKXMGWINTSTGESETF 78
DB 3 EIQLVOSGPELKKRGFTVYKISCKASGYPTNYGMNVKQAPGQGLKXMGWINTSTGESETF 62
QY 79 ADDFKRFPDSLETSANTAYLIQINNLSKSEDMATYFCARMEVYHGVYVYWGCGTTVTYVSSG 138
DB 63 AEEFKRFPDSLETSANTAYLIQINNLSKSEDMATYFCARMEVYHGVYVYWGCGTTVTYVSSG 121
QY 139 GGGSGGGGGGGSDIQLTQSHKFLSTSVGDRVSITCKASQDVNAVAVYQOKPGQSPKL 198
DB 122 SSSSGSSSSSSSDIQLTQSHKFLSTSVGDRVSITCKASQDVNAVAVYQOKPGQSPKL 181
QY 199 LIYSASSRYTGVPSRFTGSGGSPDFTFTISSVQAEIDLAVYFCQGHFRPTFTFGSGTKLEI 258
DB 182 LIYSASSRYTGVPSRFTGSGGSPDFTFTISSVQAEIDLAVYFCQGHFRPTFTFGSGTKLEI 241
QY 259 K 259

DB 242 K 242

RESULT 11

US-08-461-184-8
Sequence 8, Application US/08461184
Patent No. 5631158
GENERAL INFORMATION:
APPLICANT: DORAI, HAIMANTI
APPLICANT: OPPERMAN, HERMANN
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR HIGH PROTEIN
PRODUCTION FROM NON-NATIVE DNA
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES, INC
STREET: 45 SOUTH STREET
CITY: HOPKINTON
STATE: MA
COUNTRY: USA
ZIP: 07148
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,184
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/143,498
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: KELLEY, ROBIN D
REGISTRATION NUMBER: 34,637
REFERENCE/DOCKET NUMBER: CRP093
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/248-7000
TELEFAX: 617/248-7100
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 250 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-461-184-8

Query Match 40.2%; Score 981.5; DB 1; Length 250;
Best Local Similarity 76.8%; Pred. No. 1e-73;
Matches 185; Conservative 22; Mismatches 33; Indels 1; Gaps 1;

QY 19 QIQLVOSGPELKKRGFTVYKISCKASGYPTNYGMNVKQAPGQGLKXMGWINTSTGESETF 78
DB 3 EIQLVOSGPELKKRGFTVYKISCKASGYPTNYGMNVKQAPGQGLKXMGWINTSTGESETF 62
QY 79 ADDFKRFPDSLETSANTAYLIQINNLSKSEDMATYFCARMEVYHGVYVYWGCGTTVTYVSSG 138
DB 63 AEEFKRFPDSLETSANTAYLIQINNLSKSEDMATYFCARMEVYHGVYVYWGCGTTVTYVSSG 121
QY 139 GGGSGGGGGGGSDIQLTQSHKFLSTSVGDRVSITCKASQDVNAVAVYQOKPGQSPKL 198
DB 122 SSSSGSSSSSSSDIQLTQSHKFLSTSVGDRVSITCKASQDVNAVAVYQOKPGQSPKL 181
QY 199 LIYSASSRYTGVPSRFTGSGGSPDFTFTISSVQAEIDLAVYFCQGHFRPTFTFGSGTKLEI 258
DB 182 LIYSASSRYTGVPSRFTGSGGSPDFTFTISSVQAEIDLAVYFCQGHFRPTFTFGSGTKLEI 241
QY 259 K 259
DB 242 K 242


```

; APPLICANT: Ring, David B.
; TITLE OF INVENTION: Biosynthetic Binding Proteins For
; TITLE OF INVENTION: Imaging
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Testa, Hurwitz & Thibault/Patent Department
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,838
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Kelley, Robin D.
; REGISTRATION NUMBER: 34,637
; REFERENCE/DOCKET NUMBER: 2054/22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-248-7477
; TELEFAX: 617-248-7100
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 250 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-461-838-2

Query Match      40.2%; Score 981.5; DB 1; Length 250;
Best Local Similarity 76.8%; Pred. No. 1e-73;
Matches 185; Conservative 22; Mismatches 33; Indels 1; Gaps 1;

QY 19 QIOLVSGPELKKRGETVKISKASGYPTNYGMNVKQAPGQGLKMMGWINTSGESTF 78
Db 3 EIQLVSGPELKKRGETVKISKASGYPTNYGMNVKQAPGQGLKMMGWINTSGESTF 62
QY 79 ADFPKGRFDSLETSANTAYLQINNLSKSDMATYFCARMEVYHGVPYWGQGTIVTVSSG 138
Db 63 AEEFKGRFASLETSASTAYLQINNLSKSDMATYFCARMEVYHGVPYWGQGTIVTVSSA- 121
QY 139 GGGSGGGSGGGGSDIQLTQSHKFLSTSVGDRVSIITCKASQDVYNAVAMVYQKPGQSPKL 198
Db 122 SSSSGSSSSSSSDIQLTQSHKFLSTSVGDRVSIITCKASQDVYNAVAMVYQKPGQSPKL 181
QY 199 LIYSASRYTGVPDRFTGSGGPDFTFTISSVOAEDLAVYFCQOHRPTPTFGSGTKLEI 258
Db 182 LIYMTSRHTGVDRFTGSGGPDFTFTLTISSVOAEDLALHYCQOHRVPTPTFGSGTKLEI 241
QY 259 K 259
Db 242 K 242

RESULT 15
US-08-461-386-2
; Sequence 2, Application US/08461386
; Patent No. 5837846
; GENERAL INFORMATION:
; APPLICANT: Huston, James S.
; APPLICANT: Oppermann, Hermann
; APPLICANT: Houston, L. L.
; APPLICANT: Ring, David B.
; TITLE OF INVENTION: Biosynthetic Binding Proteins For
; TITLE OF INVENTION: Imaging
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
```

```

; ADDRESSEE: Testa, Hurwitz & Thibault/Patent Department
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,386
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Kelley, Robin D.
; REGISTRATION NUMBER: 34,637
; REFERENCE/DOCKET NUMBER: 2054/22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-248-7477
; TELEFAX: 617-248-7100
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 250 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-461-386-2

Query Match      40.2%; Score 981.5; DB 2; Length 250;
Best Local Similarity 76.8%; Pred. No. 1e-73;
Matches 185; Conservative 22; Mismatches 33; Indels 1; Gaps 1;

QY 19 QIOLVSGPELKKRGETVKISKASGYPTNYGMNVKQAPGQGLKMMGWINTSGESTF 78
Db 3 EIQLVSGPELKKRGETVKISKASGYPTNYGMNVKQAPGQGLKMMGWINTSGESTF 62
QY 79 ADFPKGRFDSLETSANTAYLQINNLSKSDMATYFCARMEVYHGVPYWGQGTIVTVSSG 138
Db 63 AEEFKGRFASLETSASTAYLQINNLSKSDMATYFCARMEVYHGVPYWGQGTIVTVSSA- 121
QY 139 GGGSGGGSGGGGSDIQLTQSHKFLSTSVGDRVSIITCKASQDVYNAVAMVYQKPGQSPKL 198
Db 122 SSSSGSSSSSSSDIQLTQSHKFLSTSVGDRVSIITCKASQDVYNAVAMVYQKPGQSPKL 181
QY 199 LIYSASRYTGVPDRFTGSGGPDFTFTISSVOAEDLAVYFCQOHRPTPTFGSGTKLEI 258
Db 182 LIYMTSRHTGVDRFTGSGGPDFTFTLTISSVOAEDLALHYCQOHRVPTPTFGSGTKLEI 241
QY 259 K 259
Db 242 K 242
```

Search completed: April 20, 2005, 07:08:35
Job time : 21.9082 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2005, 07:02:38 ; Search time 59.1549 Seconds
(without alignments)
2590.199 Million cell updates/sec

Title: US-09-596-774-6
Perfect score: 2441
Sequence: 1 MAWWTLLFLWAAKVPKQI.....LSTATKDTYDALHWQTLAPR 461

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1421835 seqs, 332370683 residues

Total number of hits satisfying chosen parameters: 1421835

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum March 0%
Maximum March 100%
Listing first 45 summaries

Database : Published Applications_AA:*

1:	/cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2:	/cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3:	/cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4:	/cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5:	/cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6:	/cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7:	/cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8:	/cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9:	/cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
10:	/cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11:	/cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12:	/cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13:	/cgn2_6/ptodata/1/pubpaa/US10_PUBCOMB.pep.*
14:	/cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15:	/cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16:	/cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
17:	/cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
18:	/cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
19:	/cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
20:	/cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	1249	51.2	631	US-10-120-1988-2
2	1015.5	41.6	443	US-10-006-773-2
3	1015.5	41.6	443	US-10-006-771A-2
4	981.5	40.2	250	US-09-887-853-2
5	969.5	39.7	250	US-10-683-547-2
6	930	38.1	444	US-08-812-393A-2
7	928	38.0	445	US-09-774-681-2
8	917.5	37.6	509	US-10-239-656-79
9	913	37.4	267	US-09-766-543-10
10	903.5	37.0	240	US-10-117-890-148
11	903.5	36.9	240	US-10-117-890-148
12	901.5	36.9	267	US-10-071-485-2
13	897	36.7	276	US-09-766-543-12

14	889.5	36.4	541	US-10-071-485-85	Sequence 85, Appl
15	888.5	36.4	711	US-10-071-485-90	Sequence 90, Appl
16	886.5	36.3	503	US-10-239-656-75	Sequence 75, Appl
17	875.5	35.9	331	US-10-059-261-169	Sequence 169, Appl
18	872.5	35.7	243	US-10-879-994-10	Sequence 10, Appl
19	872.5	35.7	243	US-10-610-452-10	Sequence 10, Appl
20	862	35.3	622	US-10-378-832A-2	Sequence 2, Appl1
21	857	35.1	248	US-09-880-748-1104	Sequence 1104, Ap
22	857	35.1	248	US-10-293-418-1104	Sequence 1104, Ap
23	854.5	35.0	251	US-09-880-748-1921	Sequence 1921, Ap
24	854.5	35.0	251	US-10-293-418-1921	Sequence 1921, Ap
25	854.5	35.0	503	US-10-239-656-77	Sequence 77, Appl
26	851	34.9	240	US-10-096-246-35	Sequence 35, Appl
27	851	34.9	240	US-10-096-246-36	Sequence 36, Appl
28	849.5	34.8	499	US-10-239-656-73	Sequence 73, Appl
29	848	34.7	240	US-10-071-485-91	Sequence 91, Appl
30	848	34.7	250	US-09-880-748-932	Sequence 932, Appl
31	846.5	34.7	249	US-10-293-418-932	Sequence 932, Appl
32	846.5	34.7	249	US-09-880-748-926	Sequence 926, Appl
33	846.5	34.7	249	US-10-293-418-926	Sequence 926, Appl
34	846	34.7	248	US-09-880-748-1446	Sequence 1446, Ap
35	846	34.7	248	US-10-293-418-1446	Sequence 1446, Ap
36	845	34.6	248	US-09-880-748-1008	Sequence 1008, Ap
37	845	34.6	248	US-09-880-748-1008	Sequence 1008, Ap
38	845	34.6	248	US-10-293-418-1008	Sequence 1008, Ap
39	845	34.6	248	US-10-293-418-1778	Sequence 1778, Ap
40	842.5	34.5	249	US-09-880-748-918	Sequence 918, Appl
41	842.5	34.5	249	US-10-293-418-918	Sequence 918, Appl
42	841.5	34.5	249	US-09-994-186-18	Sequence 18, Appl
43	841.5	34.5	249	US-10-237-667-18	Sequence 18, Appl
44	841.5	34.5	249	US-10-237-708-18	Sequence 18, Appl
45	841.5	34.5	249	US-10-237-866-18	Sequence 18, Appl

ALIGNMENTS

RESULT 1
US-10-120-1988-2
Sequence 2, Application US/101201988
Publication No. US20030215427A1
GENERAL INFORMATION:
APPLICANT: Jensen, Michael
TITLE OF INVENTION: CE7-SPECIFIC REDIRECTED IMMUNE CELLS
FILE REFERENCE: 1954-337
CURRENT APPLICATION NUMBER: US/10/120.1988
PRIOR FILING DATE: 2002-04-11
PRIORITY APPLICATION NUMBER: 60/282,859
PRIOR FILING DATE: 2001-04-11
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 631
TYPE: PRT
ORGANISM: artificial sequence
FEATURE:
OTHER INFORMATION: mouse-human chimera
US-10-120-1988-2

Query Match 51.2%; Score 1249; DB 15; Length 631;
Best Local Similarity 44.1%; Pred. No. 7.7e-79;
Matches 274; Conservative 60; Mismatches 96; Indels 192; Gaps 13;

QY	19	QIQLVSGPEIKKPGETVYKISCKASGPFITNYGMWYKQAFQGLKMGWINTSTGE	78
DB	23	QVQLQCPGAEIVKPGASVCKSGSYTFQYWMHWKQPGHLEWIGEINPNGRTNY	82
QY	79	ADDKRGFDSLETSANTAYQIINLKSEDMATYFCARWEVYHG---YVPWGGCTTYTV	135
DB	83	NRKRSKATLTVDSSSTTAFQSLGTSBDSAVYFCAR--DYITSTNFDWGGCTTLTY	140
QY	136	SSGGGSGGGGSGGGGSDIQLTQSHKFLSTSVGDRVSAITCKASQDVYNAVMYQKPGQS	195

Db 141 SSGGGGGGGGGGGSDIOMTQSSSSFSVSLGDRVITTCANEDINNRLAMVQOTPGNS 200
QY 136 PRLIYASASRTYGVPSRTGSGSGDPFTTISVQAEPLANYFCQOGRPTPTGSGTK 255
Db 201 PRLISGATNLVTVGVSRSFGSGSGKDYTLITISQAEDEFATYCCQWSTFTFGSGTE 260
QY 256 LEIK-----ALEISNVWFSSVAVLQ-- 278
Db 261 LEIKVPEKSSDKTHCPCEPAPELLGGSVFLFPKPKDMLMISTKPEVTCVVDVSHED 320
QY 279 -----KVNSTTKP-----VLR-----TPS 293
Db 321 PEVKENWYVDGVEVNAKTKPREEQYNSTYRVSVTLVHQDMLNGKEYKCKVSKALPA 380
QY 294 PVHPT---GTSQPPPE-----DCEPPGS 314
Db 361 PLEKTIKAKGQPREPOVYTLPPSRDELTKQVSLTCLVKGFPSDIAVEWESNQCPENN 440
QY 315 VKGTG--LDPLEDPKLCYLLDGLFLY----- 339
Db 441 YKTPPVLD-----SDGSFLYKSLTVDKSRMOQGNVFCGSVMHEALHNYTOKS 490
QY 340 -----GVII---TALYLKAKFSBSAETAANLQDPNOLYNELNGREE 379
Db 491 LSLSPGKMALIYLGAVAGLLFLIGLIFPRVYKFSRSDAPAYQOQNLVNEMLNGREE 550
QY 360 YDVELEKKAQREPMGKQOARRNPOEGYNNALQCKMAEAYSEIGTKERRRGKCHDLY 439
Db 551 YDVLDRKRRGRPEMGK--PRKNPOEGLYNELQCKMAEAYSEIGMKERRRGKCHDLY 609
QY 440 QGLSTATDYDALHMOTLAPR 461
Db 610 QGLSTATDYDALHMOTLAPR 631

RESULT 2

US-10-006-773-2
; Sequence 2, Application US/10006773
; Publication No. US20020132983A1
; GENERAL INFORMATION:
; APPLICANT: Jungmans, Richard P.
; TITLE OF INVENTION: Antibodies as Chimeric Effector Cell Receptors Against Tumor Anti-
; FILE REFERENCE: 003
; CURRENT APPLICATION NUMBER: US/10/006,773
; CURRENT FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: 60/250,089
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 443
; TYPE: PRT
; ORGANISM: Homo sapiens and Mus sp.
US-10-006-773-2

Query Match. 41.6%; Score 1015.5; DB 13; Length 443;
Best Local Similarity 52.0%; Pred. No. 1e-62;
Matches 225; Conservative 25; Mismatches 58; Indels 125; Gaps 8;

QY 149 GGGSDIQLTQSHKFLTSVGDVRSITCKASODVYNAVAMVQKPGSKLLIYSASRYT 208
Db 16 GVHSDIQLTQSPSSLSASVGDVITTCASQDVGSVAMVQKPGKAPKLLIYWTSTRHT 75
QY 209 GVPSPFTGSGSGDPFTTISVQAEPLANYFCQO--FRTPFTGSGTKLEIK----- 259
Db 76 GVPSPFTGSGSGDPFTTISVQAEPLANYFCQOYSLR--SFQGTKEIKRGSGSGG 132
QY 260 -----ALEISNSV----- 267
Db 133 GSGSGSGSEVQLVESGGGVVQPGSRSLRLSCSASGFPFTTYMMSWVQAQKGLIEWIGE 192
QY 268 -----MYFSS 272

Db 193 HPDSSTINVASLDRFTISRDNAKNTLFLQMSLRPEDTCVYFCASLYGFPWFAYWQ 252
QY 273 VVPLQKVNSTTKPVLRTPSPVHPGTGSP--ORPEDCPR--GSYKGTGLDFLEDPKL 328
Db 253 GTPVTVSAAKPTTTPADRPPTPA-PTIASQPLSRPEAPRPAAGAVHTRGDLPALDPKL 311
QY 329 CYLLDGLIFTYGVITLTLVLRKFSRSAETNANLQDPNOLYNELNGREEVDYLEKKA 388
Db 312 CYLLDGLIFTYGVITLTLVLRKFSRSAEPAPAYQOQNLVNEMLNGREEVDYLDKRRG 371
QY 389 RDPENGGKQOARRNPOEGYNNALQCKMAEAYSEIGTKERRRGKCHDGLYQGLSTATKD 448
Db 372 RDPENGGK--PRKNPOEGLYNELQCKMAEAYSEIGMKERRRGKCHDGLYQGLSTATKD 430
QY 449 TYDALHMOTLAPR 461
Db 431 TYDALHMOTLAPR 443

RESULT 3

US-10-006-771A-2
; Sequence 2, Application US/10006771A
; Publication No. US20020165360A1
; GENERAL INFORMATION:
; APPLICANT: Jungmans, Richard P.
; TITLE OF INVENTION: Chimeric Effector Cell Receptors Against Carcinoembryonic Antigen
; FILE REFERENCE: 002
; CURRENT APPLICATION NUMBER: US/10/006,771A
; CURRENT FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: 60/250,090
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 443
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-006-771A-2

Query Match 41.6%; Score 1015.5; DB 13; Length 443;
Best Local Similarity 52.0%; Pred. No. 1e-62;
Matches 225; Conservative 25; Mismatches 58; Indels 125; Gaps 8;

QY 149 GGGSDIQLTQSHKFLTSVGDVRSITCKASODVYNAVAMVQKPGSKLLIYSASRYT 208
Db 16 GVHSDIQLTQSPSSLSASVGDVITTCASQDVGSVAMVQKPGKAPKLLIYWTSTRHT 75
QY 209 GVPSPFTGSGSGDPFTTISVQAEPLANYFCQO--FRTPFTGSGTKLEIK----- 259
Db 76 GVPSPFTGSGSGDPFTTISVQAEPLANYFCQOYSLR--SFQGTKEIKRGSGSGG 132
QY 260 -----ALEISNSV----- 267
Db 133 GSGSGSGSEVQLVESGGGVVQPGSRSLRLSCSASGFPFTTYMMSWVQAQKGLIEWIGE 192
QY 268 -----MYFSS 272
Db 193 HPDSSTINVASLDRFTISRDNAKNTLFLQMSLRPEDTCVYFCASLYGFPWFAYWQ 252
QY 273 VVPLQKVNSTTKPVLRTPSPVHPGTGSP--ORPEDCPR--GSYKGTGLDFLEDPKL 328
Db 253 GTPVTVSAAKPTTTPADRPPTPA-PTIASQPLSRPEAPRPAAGAVHTRGDLPALDPKL 311
QY 329 CYLLDGLIFTYGVITLTLVLRKFSRSAETNANLQDPNOLYNELNGREEVDYLEKKA 388
Db 312 CYLLDGLIFTYGVITLTLVLRKFSRSAEPAPAYQOQNLVNEMLNGREEVDYLDKRRG 371
QY 389 RDPENGGKQOARRNPOEGYNNALQCKMAEAYSEIGTKERRRGKCHDGLYQGLSTATKD 448
Db 372 RDPENGGK--PRKNPOEGLYNELQCKMAEAYSEIGMKERRRGKCHDGLYQGLSTATKD 430
QY 449 TYDALHMOTLAPR 461


```

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/812.393A
  FILING DATE: 05-MAR-1997
  CLASSIFICATION: 435
  PRIOR APPLICATION DATA:
    APPLICATION NUMBER:
  FILING DATE:
  ATTORNEY/AGENT INFORMATION:
    NAME: Murashige, Kate H
    REGISTRATION NUMBER: 29,959
    REFERENCE/DOCKET NUMBER: 31333-20001.00
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: 202-887-1500
    TELEFAX: 202-822-0168
  TELEX:
  INFORMATION FOR SEQ ID NO: 2:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 444 amino acids
      TYPE: amino acid
      STRANDEDNESS: single
      TOPOLOGY: linear
    MOLECULE TYPE: protein
    FRAGMENT TYPE: internal
  US-08-812-393A-2

```

```

Query Match      38.1%; Score 930; DB 8; Length 444;
Best Local Similarity 48.2%; Pred. No. 9.4e-57;
Matches 217; Conservative 50; Mismatches 145; Indels 38; Gaps 14;

```

```

QY 21 QLVSGPE--LKKPEYTKISCKASGYPTNYGNMNVKQAPGGGLKMMGMINTSTGEFT 78
DB 24 QVQVQSPASLVLDGEENAELOCFS--IFTNQ-VQMFYORGGRLVSLLYNPGTKOS-- 78
QY 79 ADDEKGRFDFSLSEANTAYLIQINLKSMDATYFCARMEVYHGVYPMQGGTTVTYVSSG 138
DB 79 -----GRLSTTVYKERSSLSHSSQITDSGTIYCANSSGSAKLTFGKTKLSVKS 133
QY 139 GGGSGGGSGGGGSDIQLTQSHKFLSTSVGDRVSTTCASQDVYNAVAMYQKRGQSPKL 198
DB 134 GGGSGGGSGGGGSEAAVLTQSPRNKVAVTGKVTLSQNTNN-HNNMWTYRQDTGHRL 192
QY 199 LIYS---ASSRYTGVPSPRFTGS--GSGPDTFTISSVOAEIAYFC---QQHRTPTFTG 251
DB 193 IHTSYGAGSTKGDIPDGYKASRPSQENFSLLELATPSQTSVYFCASGETGNERLFFG 252
QY 252 SGTLEIKALEISNSVMYFSSVVPVLQKNSITTKKVLRTSPVPHPTGTSQPRPDCRP 311
DB 253 HGTKLSTVLT---SNSIMYFHFVFLPAKPTTT-PAHPPTPA-PTIASQPL---SLRP 304
QY 312 RGSVKGIGLDLDPKLCYLLDGLIFTYGVITLALYLRAKFSASFTANLQDPNOLYNE 371
DB 305 SSS-----RDPKLCYLLDGLIFTYGVITLALFLRVKFSRSADAPAYQGGQNLVNE 355
QY 372 LNLGRREYDVLEKKARPDEMGGKQORRRNPOEGVYNALQKDMAEAYSEIGTKGERR 431
DB 356 LNLGRREYDVLDKRGKRDPMGK-PRKNPQEGLYNELQKDMAEAYSEIGTKGERR 414
QY 432 GKGHDLGYQGLSTATKDTYDALHMQTLAPR 461
DB 415 GKGHDLGYQGLSTATKDTYDALHMQALPPR 444

```

```

RESULT 7
US-09-774-681-2
Sequence 2, Application US/09774681
Publication No. US20030208780A1
GENERAL INFORMATION:
APPLICANT: Sunol Molecular Corporation
APPLICANT: Sherman, Linda

```

```

APPLICANT: Inetgarten, Joseph
TITLE OF INVENTION: RECOMBINANT CONSTRUCTS ENCODING T CELL
FILE OF INVENTION: RECEPTORS SPECIFIC FOR HUMAN HLA-RESTRICTED TUMOR ANTIGENS
FILE REFERENCE: 31333-20001.01
CURRENT APPLICATION NUMBER: US/09/774.681
PRIOR FILING DATE: 2001-02-01
PRIOR APPLICATION NUMBER: US 08/812.393
PRIOR FILING DATE: 1997-03-05
PRIOR APPLICATION NUMBER: US 60/012,845
PRIOR FILING DATE: 1996-03-05
NUMBER OF SEQ ID NOS: 65
SOFTWARE: FASTSEQ for Windows Version 4.0
SEQ ID NO 2
LENGTH: 449
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Deduced amino acid derivative of effective T cell
US-09-774-681-2

```

```

Query Match      38.0%; Score 923; DB 10; Length 449;
Best Local Similarity 48.2%; Pred. No. 1.3e-56;
Matches 217; Conservative 50; Mismatches 145; Indels 38; Gaps 14;

```

```

QY 21 QLVSGPE--LKKPEYTKISCKASGYPTNYGNMNVKQAPGGGLKMMGMINTSTGEFT 78
DB 24 QVQVQSPASLVLDGEENAELOCFS--IFTNQ-VQMFYORGGRLVSLLYNPGTKOS-- 78
QY 79 ADDEKGRFDFSLSEANTAYLIQINLKSMDATYFCARMEVYHGVYPMQGGTTVTYVSSG 138
DB 79 -----GRLSTTVYKERSSLSHSSQITDSGTIYCANSSGSAKLTFGKTKLSVKS 133
QY 139 GGGSGGGSGGGGSDIQLTQSHKFLSTSVGDRVSTTCASQDVYNAVAMYQKRGQSPKL 198
DB 134 GGGSGGGSGGGGSEAAVLTQSPRNKVAVTGKVTLSQNTNN-HNNMWTYRQDTGHRL 192
QY 199 LIYS---ASSRYTGVPSPRFTGS--GSGPDTFTISSVOAEIAYFC---QQHRTPTFTG 251
DB 193 IHTSYGAGSTKGDIPDGYKASRPSQENFSLLELATPSQTSVYFCASGETGNERLFFG 252
QY 252 SGTLEIKALEISNSVMYFSSVVPVLQKNSITTKKVLRTSPVPHPTGTSQPRPDCRP 311
DB 253 HGTKLSTVLT---SNSIMYFHFVFLPAKPTTT-PAHPPTPA-PTIASQPL---SLRP 304
QY 312 RGSVKGIGLDLDPKLCYLLDGLIFTYGVITLALYLRAKFSASFTANLQDPNOLYNE 371
DB 305 SSS-----RDPKLCYLLDGLIFTYGVITLALFLRVKFSRSADAPAYQGGQNLVNE 355
QY 372 LNLGRREYDVLEKKARPDEMGGKQORRRNPOEGVYNALQKDMAEAYSEIGTKGERR 431
DB 356 LNLGRREYDVLDKRGKRDPMGK-PRKNPQEGLYNELQKDMAEAYSEIGTKGERR 414
QY 432 GKGHDLGYQGLSTATKDTYDALHMQTLAPR 461
DB 415 GKGHDLGYQGLSTATKDTYDALHMQALPPR 444

```

```

RESULT 8
US-10-239-656-79
Sequence 79, Application US/10239656
Publication No. US2004003839A1
GENERAL INFORMATION:
APPLICANT: KUPER, PETER
APPLICANT: RIETHMULLER, GERT
APPLICANT: LUTTERBUSE, RAUL
APPLICANT: BORSCHERT, KATRIN
APPLICANT: KISCHEL, ROMAN
APPLICANT: MAYER, MONIKA
APPLICANT: HOFMEISTER, ROBERT
TITLE OF INVENTION: MULTIFUNCTIONAL POLYPEPTIDES COMPRISING A BINDING SITE
TITLE OF INVENTION: TO AN EPTIPE OF THE NKGD2 RECEPTOR COMPLEX
FILE REFERENCE: 029976/0106

```

```

CURRENT APPLICATION NUMBER: US/10/239,656
CURRENT FILING DATE: 2003-03-06
PRIOR APPLICATION NUMBER: PCT/EP01/03144
PRIOR FILING DATE: 2001-03-26
PRIOR APPLICATION NUMBER: EP 00106467.4
PRIOR FILING DATE: 2000-03-24
NUMBER OF SEQ ID NOS: 92
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 79
LENGTH: 505
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic 3B10xP5-
US-10-239-656-79

Query Match 37.6%; Score 917.5; DB 15; Length 505;
Best Local Similarity 50.5%; Pred. No. 8,1e-56;
Matches 185; Conservative 25; Mismatches 31; Indels 125; Gaps 4

QY 19 QIQLVQSGPELKKPGEIVKISCKASGYPTNYGMNVKQAPQGLKMGMINSTGSETF 78
Db 128 QVQLQQSGPELKKPGEIVKISCKASGYPTNYGMNVKQAPQGLKMGMINSTGSEPT 187
QY 79 ADDPKGRDFSLFETSAITAYLQINNLKSEDAITFCAR----- 116
Db 188 GDDPKGRDFSLFETSAITAYLQINNLKSEDAITFCARSPDYMGQITTVVSSGGGS 247
QY 117 -----W----- 117
Db 248 EVQLLESGGGLVQPGSLKLSCKASGFDPSRYMGMWROAPKGLAEWIGEINPDSSTNY 307
QY 118 -----EYH-----GYVPYMGQITTVVSSGG 139
Db 308 TPLSKDRITIRDNANKNTLYLQMSKVRSEDTALYYCARLGQGYDYMGGITTVVSSGG 367
QY 140 GSGGGGGSGGGGGSDIQLTQSHKFLSTSVGDRVSTTCASQDVYNA-----VAMTQQKPG 193
Db 368 GSGGGGGSGGGGGSELVMTQSPPSLTITVTAAGERVTWCKSSQSLNAGNQKNYLTWYQQKPG 427
QY 194 QSPKLLIYSAASRYGVPSRFGSGSGSDPTFTTISVQAEDLAVYFCQGHFPTPTFFSG 253
Db 428 QPPLKLLIYMASTRSGVPSRFGSGSGSDPTFTTISVQAEDLAVYFCQGHFPTPTFFSG 487
QY 254 TKLEIK 259
Db 488 TKLEIK 493

RESULT 9
US-09-766-543-10
Sequence 10, Application US/09766543
Patent No. US20020041865A1
GENERAL INFORMATION:
APPLICANT: Ausetin, Richard
APPLICANT: Kwok, Cheuk S.
APPLICANT: Ring, David B.
TITLE OF INVENTION: METHODS FOR TREATING TUMORS
FILE REFERENCE: PP01679.002
CURRENT APPLICATION NUMBER: US/09/766,543
CURRENT FILING DATE: 2000-01-19
PRIOR APPLICATION NUMBER: 60/177,258
PRIOR FILING DATE: 2000-01-20
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 10
LENGTH: 267
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: 520C9
OTHER INFORMATION: humanized single-chain antibody used in the

```

```

OTHER INFORMATION: IL-2-antibody fusions
US-09-766-543-10

Query Match      37.4%; Score 913; DB 9; Length 267;
Best Local Similarity 68.4%; Pred. No. 8,1e-56;
Matches 167; Conservative 37; Mismatches 36; Indels 4; Gaps 2;

QY 19 QIQLVSGPELKKRGGEYTKISCKASGYPFTTYGNMNMVKQAQAGQGLKMMGMINTSTGESTF 78
DB 23 EMQIVESGPEPKKAGASVKVSCKASGTYFTTYGNMNMVRAQAPGQGLEWGMINTTYGOSTY 82
QY 79 ADPKRGPFDFELFESANTAYALQINNLSEDMATYFCARMEVYHGVPYMGQGTVTYVSSG 138
DB 83 ADDRERKRYMTTDTSTSTAYMDLSLSDSDDTAYTCAR---RFGFA-YMGQGLVTVSSG 138
QY 139 GGGSGGGSGGGGGSDIQLTOSHKELSVSGDRVSIITKASQSDVYNAMVYQKPGQSPKL 198
DB 139 GGGSGGGSGGGGGSDIQWTGSPSLASVSGDRVITCRASQDIGNSLTWYQKPGKTPKL 198
QY 199 LIYASASRYTVBPRFPGSGSGGPFPTTISVQMEDLAVYFCQHPKRPFFSGSTLKI 258
DB 199 LIYATSSLSIDSGVSPRFSGSGGTFTFTISSLSQPEDIATYCIQYALFPYTFGGTRLEI 258
QY 259 KALE 262
DB 259 KGSE 262

RESULT 10
US-10-127-890-148
Sequence 148, Application US/10127890
Publication No. US20030166196A1
GENERAL INFORMATION:
APPLICANT: Betcer, Marc D.
Carroll, Stephen F.
Studnicka, Gary M.
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
Proteins
NUMBER OF SEQUENCES: 173
CORRESPONDENCE ADDRESS:
ADDRESSEE: McAndrews, Held & Malloy, Ltd.
STREET: 500 West Madison Street, 34th floor
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60661
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/127,890
FILING DATE: 23-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/646,360
FILING DATE: 13-MAY-1996
APPLICATION NUMBER: PCT/US94/05348
FILING DATE: 12-MAY-1994
APPLICATION NUMBER: US 08/064,691
FILING DATE: 12-MAY-1993
APPLICATION NUMBER: US 07/988,430
FILING DATE: 09-DEC-1992
APPLICATION NUMBER: US 07/901,707
FILING DATE: 19-JUN-1992
APPLICATION NUMBER: US 07/787,567
FILING DATE: 04-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: McNicholas, Janet M.
REGISTRATION NUMBER: 32,918
REFERENCE/DOCKET NUMBER: 200-70-P4
TELECOMMUNICATION INFORMATION:

```


FEATURE:
OTHER INFORMATION: Synthetic
US-10-071-485-2

Query Match 36.9%; Score 901.5; DB 14; Length 267;
Best Local Similarity 68.3%; Pred. No. 5.2e-55;
Matches 172; Conservative 28; Mismatches 49; Indels 3; Gaps 2;

QY 8 LFLMAAKVPOKIOIVQSGPELKPKPEYKISCKASGYFTNYGMNVQAPQGLKMMG 67
DB 12 LLLLAOPMAQVOLVQSGSELKPKQASVKISCKASGYFTIDYGMNVQAPQGLKMMG 71
QY 68 WINTSGESTFADPKGRFDSLETSAANTAYLOINNLKSEDMATYFCARMEYHGVYRW 127
DB 72 WINTYTGESTYVDFKGRFVSLDTSVSAAYLOISLKAEDTATYFCARGFY--AMDYN 129
QY 128 GGGTTVTVSSGGGGGGGGGGSDIOLTQSHKFLSTSVGRVSTICKASODVYNAVAM 187
DB 130 GGGTTVTVSSGGGGGGGGGGSDIOLTQSPATMASPGRVTLTCSASSI-SYMFV 188
QY 168 YQKKGQSPKLITYASRVTGVPSPFTGSGSPDFTTISVQADLAVYFCQGHFRTP 247
DB 169 YHQRQGSRRLLIYDTSNLASGVPAFSGSGSTYSILTISHMEPDFTATYFCOSSSY 248
QY 248 FTFGSGTKLEIK 259
DB 249 FTFGSGTKLEIK 260

RESULT 13

US-09-766-543-12
Sequence 12, Application US/09766543
Patent No. US20020041865A1

GENERAL INFORMATION:
APPLICANT: Austin, Richard
APPLICANT: Kwok, Cheuk S.
APPLICANT: Ring, David B.
TITLE OF INVENTION: METHODS FOR TREATING TUMORS
FILE REFERENCE: P01679.002
CURRENT APPLICATION NUMBER: US/09/766,543
CURRENT FILING DATE: 2000-01-19
PRIOR APPLICATION NUMBER: 60/177,258
PRIOR FILING DATE: 2000-01-20
NUMBER OF SEQ ID NOS: 14
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO 12
LENGTH: 276
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: H520C9aFv plus
US-09-766-543-12

Query Match 36.7%; Score 897; DB 9; Length 276;
Best Local Similarity 69.7%; Pred. No. 1.1e-54;
Matches 168; Conservative 34; Mismatches 35; Indels 4; Gaps 2;

QY 19 QIOLVQSGPELKPKPEYKISCKASGYFTNYGMNVQAPQGLKMMGINTSGESTF 78
DB 23 EIQLVQSGPEVKKPGASVKISCKASGYTFANYGMNMVQAPQGLKMMGINTYGSSTY 82
QY 79 ADDFKGRFDSLETSAANTAYLOINNLKSEDMATYFCARMEYHGVYRWGGTTTVSSG 138
DB 83 ADDFKGRFDSLETSAANTAYLOINNLKSEDMATYFCARMEYHGVYRWGGTTTVSSG 138
QY 139 GGGSGGGGGGGGGSDIOLTQSHKFLSTSVGRVSTICKASODVYNAVAMVQKPGOSPRL 198
DB 139 GGGSGGGGGGGGGSDIOLTQSHKFLSTSVGRVSTICKASODVYNAVAMVQKPGOSPRL 198
QY 199 LIYASASRYTVSPSRFTGSGSPDFTTISVQADLAVYFCQGHFRTPFTFGSGTKLEI 258
DB 199 LIYATSSIDSQVSPSRFTGSGSPDFTTISVQADLAVYFCQGHFRTPFTFGSGTKLEI 258

QY 259 K 259
DB 259 K 259

RESULT 14

US-10-071-485-85
Sequence 85, Application US/10071485
Publication No. US20030099648A1

GENERAL INFORMATION:
APPLICANT: Buysse, Marie-Ange
APPLICANT: Sablon, Erwin
TITLE OF INVENTION: INTERFERON-gamma-BINDING MOLECULES FOR TREATING SEPTIC
TITLE OF INVENTION: SHOCK,
TITLE OF INVENTION: CACHEXIA, IMMUNE DISEASES AND SKIN DISORDERS
FILE REFERENCE: INNS:015
CURRENT APPLICATION NUMBER: US/10/071,485
CURRENT FILING DATE: 2002-02-07
PRIOR APPLICATION NUMBER: 09/485,737
PRIOR FILING DATE: 2000-02-14
PRIOR APPLICATION NUMBER: PCT/EP 98/05165
PRIOR FILING DATE: 1998-08-14
PRIOR APPLICATION NUMBER: EPO 96870139.7
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: EPO 97870122.5
NUMBER OF SEQ ID NOS: 104
SOFTWARE: Patentin version 3.0
SEQ ID NO 85
LENGTH: 541
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: SYNTHETIC
US-10-071-485-85

Query Match 36.4%; Score 889.5; DB 14; Length 541;
Best Local Similarity 67.7%; Pred. No. 7.9e-54;
Matches 168; Conservative 30; Mismatches 47; Indels 3; Gaps 2;

QY 19 QIOLVQSGPELKPKPEYKISCKASGYFTNYGMNVQAPQGLKMMGINTSGESTF 78
DB 1 QIOLVQSGPELKPKQASVKISCKASGYFTIDYGMNVQAPQGLKMMGINTYGSSTY 60
QY 79 ADDFKGRFDSLETSAANTAYLOINNLKSEDMATYFCARMEYHGVYRWGGTTTVSSG 138
DB 61 VDDFKGRFVSLDTSVSAAYLOISLKAEDTATYFCARGFY--AMDYWGQTTTVSSG 118
QY 139 GGGSGGGGGGGGGSDIOLTQSHKFLSTSVGRVSTICKASODVYNAVAMVQKPGOSPRL 198
DB 119 GGGSGGGGGGGGGSDIOLTQSPATMASPGRVTLTCSASSI-SYMFVYHQRQGSRRLL 177
QY 199 LIYASASRYTVSPSRFTGSGSPDFTTISVQADLAVYFCQGHFRTPFTFGSGTKLEI 258
DB 178 LIYDTSNLASGVPAFSGSGSTYSILTISHMEPDFTATYFCOSSSYFTFGSGTKLEI 237
QY 259 KALEISNS 266
DB 238 KRTPLGDT 245

RESULT 15
US-10-071-485-90
Sequence 90, Application US/10071485
Publication No. US20030099648A1
GENERAL INFORMATION:
APPLICANT: Buysse, Marie-Ange
APPLICANT: Sablon, Erwin
TITLE OF INVENTION: INTERFERON-gamma-BINDING MOLECULES FOR TREATING SEPTIC
TITLE OF INVENTION: SHOCK,
TITLE OF INVENTION: CACHEXIA, IMMUNE DISEASES AND SKIN DISORDERS
FILE REFERENCE: INNS:015

```

; CURRENT APPLICATION NUMBER: US/10/071,485
; CURRENT FILING DATE: 2002-02-07
; PRIOR APPLICATION NUMBER: 09/485,737
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: PCT/EP 98/05165
; PRIOR FILING DATE: 1998-08-14
; PRIOR APPLICATION NUMBER: EPO 98870139.7
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: EPO 97870122.5
; PRIOR FILING DATE: 1997-08-18
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 90
; LENGTH: 711
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: SYNTHETIC
; US-10-071-485-90

```

```

Query Match          36.4%; Score 888.5; DB 14; Length 711;
Best Local Similarity 69.7%; Pred. No. 1.3e-53;
Matches 168; Conservative 27; Mismatches 43; Indels 3; Gaps 2;

```

```

QY 19 QIQVQSGPELKKPGETVYKISCKASGYPTNYGNMNVKQAPGQGLKMMGWINTSTGESTF 78
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 473 OVQVQSGSELKKPGASVYKISCKASGYPTDGYMNMVVKQAPGQGLKMMGWINTYTGESTY 532
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 79 ADDFKGRFDFSLSTANTAYLQINNKKSEDMATYFCARMEVYHGVVPYMGQGTVTYVSSG 138
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 533 VDFKGRFVFLSDIVSNAIYLQISLKAEDTATYFCARRGFY--ANDYMGQGTVTYVSSG 590
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 139 GGGSGGGSGGGSDIQLTQSHKFLSTSVGDRVSITCKASQDVYNVAVMYOQKPGQSPKL 198
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 591 GGGSGGGSGGGSDIVLTQSPATWSASGERVTLTCSASSI--SYMFMYHQRPQSPRL 649
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 199 LIYSASSRYTVPSKFTSGSGSPDFTTISVQAEIDLAVYFCQCHFRTPFTFGSGTKLEI 258
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 650 LIYDTSNLASGVPARFSGSGSTSYLTI SRMEPEDFATYFCHOSSSYPTFGQGTKLEI 709
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY .259 K 259
   |
DB 710 K 710

```

Search completed: April 20, 2005, 07:10:38
 Job time : 61.1549 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 20, 2005, 10:01:29 ; Search time 6166.38 Seconds
(without alignments)
3622.520 Million cell updates/sec

Title: US-09-596-774-6

Perfect score: 2441
Sequence: 1 MAMWTLFLMAAAKVPKQI.....ISTATKDTYDALHMQTLAPR 461

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 2422767955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_plus_p2n.model -DEV=xlh
-O=/cgn2.1/USPTO_epool_h/US05956774/runcat_20042005_075442_12508/app_query.fasta_1.1230
-DB=genEmbl -QMT=faeap -SUFFIX=ap20.1.rge -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsu62 -TRANS=human4.0.cdi -LIST=45
-DOCLN=200 -THR SCORE=100 -THR MAX=100 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEPRTSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US05956774 @CGN_1_1.663 @runcat_20042005_075442_12508 -NCPU=6 -ICPU=3
-NO MMAP -LARGEOQUERY -NEG SCORES=0 -WAIT -DSPBLCK=100 -LONGCLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:
1: gb_da:
2: gb_hcg:
3: gb_in:
4: gb_om:
5: gb_ov:
6: gb_pat:
7: gb_ph:
8: gb_pl:
9: gb_pr:
10: gb_ro:
11: gb_ers:
12: gb_by:
13: gb_un:
14: gb_vl:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2441	100.0	1479	6	A47281 Sequence 5
2	1506.5	61.7	3058	6	A93936 Sequence 1
3	1503.5	61.6	1284	6	A93938 Sequence 3
4	1315.5	53.9	2012	6	A22539 M.musculus/

5	1315.5	53.9	2012	6	128526	128526 Sequence 13
6	1312.5	53.8	2070	6	A57337	A57337 Sequence 7
7	1312.5	53.8	2730	6	A57353	A57353 Sequence 23
8	1308	53.6	894	12	XXU64991	U64994 Synthetic C
9	1308	53.6	1862	6	A50995	A50995 Sequence 36
10	1308	53.6	1862	6	AR268367	AR268367 Sequence
11	1308	53.6	1919	6	A50993	A50993 Sequence 34
12	1308	53.6	1919	6	AR268366	AR268366 Sequence
13	1308	53.6	2214	6	A57341	A57341 Sequence 11
14	1308	53.6	2793	6	A57343	A57343 Sequence 13
15	1308	53.6	2793	6	A57351	A57351 Sequence 21
16	1308	53.6	3177	6	A57347	A57347 Sequence 17
17	1305	53.5	1692	6	A50960	A50960 Sequence 1
18	1305	53.5	1692	6	AR268336	AR268336 Sequence
19	1303	53.4	2233	6	A22470	A22470 M.musculus/
20	1303	53.4	2233	6	128521	128521 Sequence 6
21	1302	53.3	794	6	A57333	A57333 Sequence 3
22	1302	53.3	2718	6	A57349	A57349 Sequence 19
23	1302	53.3	2793	6	A57345	A57345 Sequence 15
24	1285.5	52.7	748	6	A22469	A22469 M.musculus
25	1285.5	52.7	748	6	A47277	A47277 Sequence 1
26	1285.5	52.7	748	6	128520	128520 Sequence 4
27	1013	41.5	726	6	AX100174	AX100174 Sequence
28	1011.5	41.4	1422	6	A63768	A63768 Sequence 32
29	1011.5	41.4	1545	6	A63770	A63770 Sequence 34
30	994	40.7	771	6	AX057978	AX057978 Sequence
31	981.5	40.2	909	6	AR007979	AR007979 Sequence
32	981.5	40.2	909	6	AR058994	AR058994 Sequence
33	981.5	40.2	909	6	123444	123444 Sequence 1
34	981.5	40.2	909	6	143364	143364 Sequence 7
35	981.5	40.2	909	6	161419	161419 Sequence 1
36	981.5	40.2	909	6	196033	196033 Sequence 7
37	978.5	40.1	987	6	AX935453	AX935453 Sequence
38	975.5	40.0	777	6	AX935402	AX935402 Sequence
39	963.5	39.5	987	6	A935455	A935455 Sequence
40	962.5	39.4	1956	6	A63774	A63774 Sequence 38
41	962.5	39.4	2079	6	A63776	A63776 Sequence 40
42	960.5	39.3	777	6	AX935404	AX935404 Sequence
43	942	38.6	1632	12	AY452127	AY452127 Synthetic
44	942	38.6	1761	12	AY452134	AY452134 Synthetic
45	942	38.6	3195	12	AY452128	AY452128 Synthetic

ALIGNMENTS

RESULT 1
A47281 LOCUS A47281 1479 bp DNA linear PAT 07-MAR-1997
DEFINITION Sequence 5 from Patent WO9530014.
ACCESSION A47281
VERSION A47281.1 GI:2301300
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 1479)
AUTHORS Groner, B. and Moritz, D.
TITLE BIFUNCTIONAL PROTEIN, PREPARATION AND USE
JOURNAL Patent: WO 9530014-A 5 09-NOV-1995;
CIBA GEIGY AG (CH)
COMMENT Other publication ZA 9503440 951102
Other publication AU 2446995 951129.

FEATURES
source
1..1479
Location/Qualifiers

/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
40..1425
/codon_start=1
/product="LEADER-SCPV (FRP5):LYT-2 HINGE-ZETA"
/protein_id="CAA03050.1"
/db_xref="GI:2301301"
/translation="MAMWTLFLMAAAKVPKQIOLVOSGPBLKKPBTIVISKASG"

YPTNYGNMVKQAPQGLKMMGINTSGESTPADDEKFRPFSLETSSANAYLOIN
 NKSEDMATYPCARMEVTHGVYIPYNGOCTTVTVSSGGSSGGSSGGSSGGSDLOLQSH
 KRLSTVGRVSVITCKASQDPAVNAVAMTQKFGSGKLEITASSRTYGVSPFTGSG
 SCDFPTFTISSVQADBLAVYPCQOHFRTPFTGSGKLEIKALEISNMYFSSVVP
 LQKVSSTTPKPLRTSPVHPPTGSPQOPBEDCRPNRGVTKGLDFLEBPKCYLLDG
 ILFIVGIIITALLYRAKFSRASAETAMLDPPQLYENLNGRREEDVLEKRRADPE
 MGKQOORRNPOEGVYVALQKDXMAEAYSEICTKGBRRRKGHDGLYQGLSATYMDTY
 DALHMQTLAPR"

ORIGIN

Alignment Scores:

Pred. No.: 1.18e-180 Length: 1479
 Score: 2441.00 Matches: 461
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 Gaps: 0

US-09-596-774-6 (1-461) X A47281 (1-1479)

QY 1 MetAlaTriPValTriPThrLeuLeuPheLeuMetAlaAlaAlaValProlyGlnIle 20
 Db 40 ATGGCTTGGGTGATGACCTTCTTCTGATGGACGCTGCCAAGTGCCCAAGCAGATC 99
 QY 21 GlnLeuValGlnSerGlyProGlnLeuYslyProGlyGlnThrValIleSerCys 40
 Db 100 CAGTTGGTGACGTGAGACCTGAGCTGAAGAAGCCTGGAGAGACAGTCAAGTCTCTGC 159
 QY 41 LysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTriPValIleGlnAlaProGly 60
 Db 160 AAGGCTCTGGGTATCTTTCACAACTATGAAATGAACTGGGTCAACAGGCTTCACAGA 219
 QY 61 GlnGlyLeuYsTriPMeGlyTyrIleAsnThrSerThrGlySerThrPheAlaAsp 80
 Db 220 CAGGCTTTAAGTGAATGGGCTGAGATTAAACCTTCACTGAGAGATCAACATTTGCTCAT 279
 QY 81 AspPheYsGlyArgPheAspPheSerLeuGlnThrSerAlaAsnThrAlaTyrLeuGln 100
 Db 280 GACTTCAGAGGACGGTGTGACTTCTTGTGMAAAGCTCGCCAAACACTGCTATTGTCAG 339
 QY 101 IleAsnLeuYsSerGlnAspMetAlaThrTyrPheCysAlaArgTriPValIleTyr 120
 Db 340 ATCAACAACCTCAAAAGTGAAGACATGCTACATATTTCTGTGCAAGATGAGAGGTTTAC 399
 QY 121 HisGlyTyrValProTyrTriPValGlnGlyThrThrValThrValSerSerGlyGlyGly 140
 Db 400 CACGGCTACGTTCTTCTTACCTGGGCGCAAGACACGCTTCTTCTTCTGCGGCTGCG 459
 QY 141 GlySerGlyGlyGlySerGlyGlyGlyGlySerAspIleGlnLeuThrGlnSerHis 160
 Db 460 GGTCTGTGGTGGCGGTGCGGTGCGGTGCGGTCTGTGACATCCAGCTGACCCAGCTCTCAC 519
 QY 161 LysPheLeuSerThrSerValGlyAspArgValSerIleThrCysIleValSerGlnAsp 180
 Db 520 AAATTTCTCTGCTCACTTCACTGAGAGACAGGCTCAGCATCACTGCAAGGCCAGTCAGAT 579
 QY 181 ValIleAsnAlaValAlaTriPValGlnGlnYsProGlyGlnSerProlySLeuLeuIle 200
 Db 580 GTGATATATGCTGTGCTGCTGATCAACAGAAACAGAGACATCTCTAACTCTGATT 639
 QY 201 TyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGlySerGly 220
 Db 640 TACTGGGATCTCTCCGGTACACAGATGCTCTTCTCTCTTCACTGCGAGCTCTGGG 699
 QY 221 ProAspPheThrPheThrIleSerSerValGlnAlaGlnAspLeuAlaValIlePheCys 240
 Db 700 CCGGATTTCACTTTCACATCAGCAGTGTGAGGCTGAAGCTCTGGCAGTTTATTCTGT 759
 QY 241 GlnGlnHisPheArgThrProPheThrPheGlySerGlyThrIleYsLeuGlnIleYsAla 260
 Db 760 CAGCAACATTTTCTTACCTCATTCACTGCGGTGCGGGAACAAATGAGAGATCAAGCT 819
 QY 261 LeuGlnIleSerAsnSerValMetTyrPheSerSerValValProValLeuGlnIleYsVal 280

Db 820 CTAGAGATCAACAACCTGGTATGATCTACTGCTGTCTGCGCAAGTCTTTCAGAAAGTG 879
 QY 281 AsnSerThrThrThrIleYsProValLeuArgThrProSerProValHisProThrGlyThr 300
 Db 880 AACTCTACTACTTACCAAGCCAGTCTGGAACCTCTTCACTGTGACACCTTACCGGAGCA 939
 QY 301 SerGlnProGlnAlaArgProGlnAspCysArgProArgGlySerValIleYsGlyThrGlyLeu 320
 Db 940 TCTCAGCCCCAGAGACCAAGAAATGTCGCGCCGTGCTCAGTGAAGGGGACCGGATTTG 999
 QY 321 AspPheLeuGlnAspProlySLeuCysTyrLeuLeuAspGlyIleLeuPheIleTyrGly 340
 Db 1000 GACTTCTTCAAGGATCCCAAACTCTGCTACTTGTCAATGGAATCTTCTTCAATCTACGA 1059
 QY 341 ValIleIleThrAlaLeuTyrLeuArgAlaIleYsPheSerArgSerAlaGlnThrAlaAla 360
 Db 1060 GTCATCATCAACAGCCCTGTACTGAGAGCAAAATTCAGCAGAGAGTGCAGAGACTGTGCC 1119
 QY 361 AsnLeuGlnAspProAsnGlnLeuTyrArgGlnLeuAsnLeuGlyArgArgGlnGlnIleTyr 380
 Db 1120 AACTCGAGGACCCCAACAGCTTCAATGAGCTCAATCTAGGGCGAAGAGAAATAT 1179
 QY 381 AspValLeuGlnIleYsArgAlaArgAspProGlnMetGlyIleYsGlnGlnIleArgArg 400
 Db 1180 GAGCTCTTGGAGAAAGAGCGGGCTCGGATCCAGAGATGGAGACAAACAGCAGAGAGAG 1239
 QY 401 ArgAsnProGlnGlnGlyValIleTyrAsnAlaLeuGlnIleYsAspYsMetAlaGlnAlaTyr 420
 Db 1240 AGGAACCCCAAGAGAGCGTATCAATGCAATGCACTGACAGAAACAAATGCGAGAGCCTAC 1299
 QY 421 SerGlnIleGlyThrIleYsGlyGlyIleArgArgArgGlyIleYsGlyHisAspGlyLeuTyrGln 440
 Db 1300 AGTCAGATCCGACAAAGCGGAGAGCGGAGAGCGCAAGGGGCAAGTGCCCTTTACAG 1359
 QY 441 GlyLeuSerThrAlaThrIleYsAspThrTyrAspAlaLeuHisMetGlnThrLeuAlaPro 460
 Db 1360 GGTCTGAGCAGCTGCAACAGACACTTGTATGATGCCCTGATATGACAGACCTTGGCCCT 1419
 QY 461 Arg 461
 Db 1420 CGC 1422

RESULT 2
 A93936 3058 bp DNA linear PAT 22-JAN-2000
 LOCUS
 DEFINITION Sequence 1 from Patent WO9720938.
 ACCESSION A93936
 VERSION A93936.1 GI:6742038
 KEYWORDS
 SOURCE unidentified
 ORGANISM unidentified
 REFERENCE 1 (bases 1 to 3058)
 AUTHORS Rosenthal, F. and Kulmburg, P.
 TITLE CELLS WITH HYBRID RECEPTOR AND GENE CONSTRUCT WHICH CAN BE
 CONTROLLED BY SAID HYBRID RECEPTOR, AND USE OF SAID CELLS IN GENE
 THERAPY
 JOURNAL Patent: WO 9720938-A.1 12-JUN-1997.
 FEATUERS location/Qualifiers
 source 1..3058
 /organism="unidentified"
 /mol_type="unassigned DNA"
 /db_xref="taxon:32644"

ORIGIN

Alignment Scores:
 Pred. No.: 1.76e-107 Length: 3058
 Score: 1506.50 Matches: 299
 Percent Similarity: 75.27% Conservative: 45
 Best Local Similarity: 65.43% Mismatches: 86
 Query Match: 61.72% Indels: 27

DB: 6 Gaps: 6

US-09-596-774-6 (1-461) x A93936 (1-3058)

9 PheleuMetAlaAlaAlaValProlysgInilegInleuValGlnSerGlyProGlu 28

832 TTTGCTTTCTCTCCACAGAGTGTCCACTCCCGCTCAACTGCACAGCTGGGGCTGAG 891

29 LeuylsValProGlyGluThrValylsSerCylsValAsnSerGlyTyrProPheThr 48

892 CTTGGAAGCCCTGGGGCTTCAAGTGAAGCTGCTCCAGAGCTTCTGGCTACCTTACC 951

49 AsnTyrGlyMetAsnTyrValylsGlnAlaProGlyGlnGlyLeuylsTyrPheMetGlyTyr 68

952 AGCTACTGATGCTGCTGGGTGAAGAGGCTTGACAGAGGCTTGAGTGAATGGAGG 1011

69 IleAsnThrSerThrGlyGluSerThrPheAlaAspPhePheLysGlyYargPheAspPhe 88

1012 ATTGATCTTAATAGTGTGTACTAAGTACAAATGGAAGTTCAAGAGCAAGCCACACTG 1071

89 SerLeuGluThrSerAlaAsnThrAlaThrValylsLeuGlnIleAsnAsnLeuLysSerGluAsp 108

1072 ACTGTAGACAAACCTCTCAGCAGCAGCTACATGACCTCAGAGCTGACTGAGAGC 1131

109 MetAlaThrTyrPheCysAlaAlaArgTyrGluValTyr---HisGlyTyrValProTyrTyr 127

1132 TCTGGGCTCTATTATTGTGCAGAGATACATTAATCTACGAGTACTTGTACTACTG 1191

128 GlyGlnGlyThrThrValThrValSerSerGlyGlyGlySerGlyGlyGlySer 147

1192 GGCCAAAGGAGACCAAGCTCAGCTCTCTCAGGTGAGGGGGGTTGAGGGAGTGGCTCT 1251

148 GlyGlyGlyGlySerAspIleGlnLeuThrGlnSerHisLysPheLeuSerThrSerVal 167

1252 GGGGTGGCGGATCTCAGCGTGTGGAGACAG---GAATCTGCACCTCACATCACCCT 1308

168 GlyAspAspValSerIleThrCysAlaValSerGlnAspValTyr-----AsnAla 184

1309 GGTGAACAGTACACCTACTGTCTGCTCAATACTGGGGCTGTACCAACTAGTAACTAT 1368

185 ValAlaTyrTyrGlnGlnLysProGlyGlnSerProLysLeuLysIleTyrSerAlaSer 204

1369 GCCAACTGGGTCCAGAAACACAGATCATTTATTCACTGCTGTAATAGTGTACCAAC 1428

205 SerArgTyrThrGlyValProSerArgPheThrGlySerGlySerGlyProAspPheThr 224

1429 AACCAGCTCCAGGTGTTCTGCTCAGATTCTCAGGCTCCCTGATTGAGACAAAGCTGCC 1488

225 PheThrIleSerSerValGlnAlaGlnAspLeuAlaValTyrPheCysGlnGlnHisPhe 244

1489 CTCACCATCACAGGGGACAGACTGAGTGAAGGCAATATATTTCTGTGCTTATGCTAC 1548

245 ArgThrProPheThrPheGlySerGlyThrLysLeuGlnIleLysAlaLeuGlnLysSer 264

1549 AGCAACACTTGGGTGTTGGTGAAGAACCAACGACGCTGCTCAGAG-----CTCAG- 1596

265 AsnSerValMetTyrPheSerSerValValProValLeuGlnLysValAsnSerThrThr 284

1597 -----GAATTCACCTACT 1608

285 ThrLysProValLeuAlaArgThrProSerProValHisProThrGlyThrSerGlnProGln 304

1609 ACCAAGCCAGTGTCTGCAACTCTGCTCACTGTGCACCTTACCGGACATCTACGCCCCAG 1668

305 ArgProGlnAspCysArgProArgGlySerValylsGlyTyrGlyLeuAspPheLeuGln 324

1669 AGACCAAGAAATGTGCGGCCCTGCTCAGTGAAGGGGACCGGATTGGACTTC----- 1722

325 AspProLysLeuCysTyrIleLeuLysAspGlyIleLeuPheIleTyrGlyValIleIleThr 344

1723 GACCCCAAACTCTGTACTTGTAGATGAATCCCTTCTATCTACGAGATCTCATCTACA 1782

345 AlaLeuTyrLysLeuArgAlaLysPheSerArgSerAlaGluThrAlaAlaLeuGlnAsp 364

DB 1783 GCCCTGACTGACGCGAAATTCAGCAGAGTGCAGAGACTGCTCCAACTCGCAGGAC 1842

365 ProAsnGlnLeuTyrAsnGlnLeuAsnLeuGlnYargAspGlnGluTyrAspValLeuGln 384

1843 CCGAACAGGCTCTCAATGAGCTCAATGTAGGGCGAAGAGGAATATGACCTTTGGAG 1902

385 LysLysArgAlaAspAspProGlnMetGlyLysGlnGlnIleArgArgAsnProGln 404

1903 AAGAAAGCGGCTCGGATCCAGAGATGGGAGCAAAACAGCAGAGAGAGAAACCCCGAG 1962

405 GluGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaGluAlaTyrSerGluIleGly 424

1963 GAAGCATATCAATATGCACTGACAGAAACAAATGCGAGAAAGCTTACAGTGAATCGGC 2022

425 ThrLysGlyGluArgArgGlyLysGlyHisAspGlyLeuTyrGlnGlyLeuSerThr 444

2023 ACMAAGCGAGAGCGGAGAGGAGGAGGAGGAGCATGTGCTTTACAGAGGCTTCAGCACT 2082

445 AlaThrLysAspThrTyrAspAlaLeuHisMetGlnThrLeuAlaProArg 461

2083 GCCACMAAGACACCTATGATGCTCTGCATATGCAAGCTGGCCCTCGC 2133

RESULT 3

A93938 1284 bp DNA linear PAT 22-JAN-2000

LOCUS Sequence 3 from Patent WO9720938.

DEFINITION A93938

ACCESSION A93938.1 GI:6742040

VERSION

KEYWORDS

SOURCE

ORGANISM

unidentified

unclassified.

REFERENCE

1 (bases 1 to 1284)

ROSENTHAL, F. and KULMBURG, P.

CELLS WITH HYBRID RECEPTOR AND GENE CONSTRUCT WHICH CAN BE

CONTROLLED BY SAID HYBRID RECEPTOR, AND USE OF SAID CELLS IN GENE

THERAPY

JOURNAL Patent: WO 9720938-A 3 12-JUN-1997;

ROSENTHAL, FELICIA (DE); KULMBURG, PETER (DE)

FEATURES

Location/Qualifiers

1..1284

source

/organism="unidentified"

/mol_type="unassigned DNA"

/db_xref="taxon:32644"

ORIGIN

Alignment Scores:

Pred. No.: 1,01e-107 Length: 1284

Score: 1503.50 Matches: 298

Percent Similarity: 76.00% Conservative: 44

Best Local Similarity: 66.22% Mismatches: 81

Query Match: 61.59% Indels: 27

DB: 6 Gaps: 6

US-09-596-774-6 (1-461) x A93936 (1-1284)

16 ValProlysgInilegInleuValGlnSerGlyProGluLeuLysLysProGlyGluThr 35

4 GTCCACTCCAGGCTCCAACTGACAGCAGTGGGGCTGAGCTTGTGTAACCTGGGGCTTCA 63

36 ValylsIleSerCylsValAsnSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrVal 55

64 GTGAAGCTGTCTGCAAGGCTTCTGGCTACACCTTACCAAGCTACTGATGACTGGGTG 123

56 LysGlnAlaProGlyGlnGlyLeuylsTyrPheMetGlyTyrIleAsnThrSerThrGlyGlu 75

124 AAGCAGAGGCTTGAGCAGAGGCTTGAGTGAAGATGATCCAAATAGTGTGCT 183

76 SerThrPheAlaAspAspPheLysGlyYargPheAspPheSerLeuGluThrSerAlaAsn 95

184 ACTAAGTACAAATGAGAGTTCAGAGCAAGGCACACTGACTGTAGAGCAAAACCTTCAGC 243

```

Oy 96 ThrAlaTYrLeuGlnIleAsnAsnLeuLYSerGluAspMetAlaThrTYrPheCysAla 115
Db 244 AACGCTTACATGAGCTGACAGCCGACATCGAGACTCTGGCGTCTATTATTGTGCA 303
Oy 116 ArgTYrGluValTYr---HisGlyTYrValProTYrTYrGlnGlnIYThrThrValThr 134
Db 304 AGATACGATTACTACGGGTAGTAGTACTTGTGACTATGGGGCCAAAGGACCGAGCTCAC 363
Oy 135 ValSerSerGlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlySeraspIle 154
Db 364 GTCCTCTCAGGTGGAGCGGCTTCAGCGGAGGTGCTCGCGGTGGCGGATCTCAGGCT 423
Oy 155 GlnLeuThrGlnSerHisLYSPhelSerThrSerValGlyAspArgValSerIleThr 174
Db 424 GTTGGGACACAG---GAATCTGCATCTACACACATCCTGGTGAACAGTCACTCACT 480
Oy 175 CysLYsAlaSerGlnAspValTYr-----AsnAlaValAlaTYrTYrGlnGlnIYs 191
Db 481 TGTGCTCAAGTACGTGGGCTGTACAACACTAGTACCTAGCCAACTGGGTCCAAAGAAA 540
Oy 192 ProGlyGlnSerProLYsLeuLeuIleTYrSerAlaSerSerArgTYrThrGlyValPro 211
Db 541 CCAGATCATTTATTACATGCTATAGTGTGTACCAACACGAGCTCCAGGTTCTTCT 600
Oy 212 SerArgPheThrGlySerGlySerGlyProAspPheThrPheThrIleSerSerValGln 231
Db 601 GCCAATTTCTCAGGCTCTCCATGAGACAGAGCTCCCTCACCATTACAGGGGCACAG 660
Oy 232 AlaGluAspLeuAlaValTYrPheCysGlnGlnHisPheArgThrProPheThrPheGly 251
Db 661 ACTGAGGATGAGGCAATATATTTCTGTCTCATGTGTCACACCAACCATTTGGTTCGCT 720
Oy 252 SerGlyThrLYsLeuGlnIleLYsAlaLeuGlnIleSerAsnSerValMetTYrPheSer 271
Db 721 GGAGAACCAAACTACACTGTC-----CTCAG----- 747
Oy 272 SerValValProValIleuGlnLYsValAsnSerThrThrThrLYsProValIleuArgThr 291
Db 748 -----GAATTCATCTACCAAGCCGCTCTCGAACT 780
Oy 292 ProSerProValHisProThrGlyThiSerGlnProGlnIYAspCysArgPro 311
Db 781 CCCCTACCTGTGACCCCTACCGGACATCTCAGCCCCAGAGACCAAGATTTGCGGCC 840
Oy 312 ArgGlySerValLYsGlyTYrGlyLeuAspPheLeuGluAspProLYsLeuCysTYrIleu 331
Db 841 CTGGCTCAGTGAAGGGAGCCGAGTTGACCTTC-----GACCCCAAACTCTGCTACTTG 894
Oy 332 LeuAspGlyIleLeuPheIleTYrGlyValIleIleThrAlaLeuTYrLeuArgAlaLYs 351
Db 895 CTAGATGGAATCCTCTTCATCTACGAGTCAATCAACGCCCTGTACTGAGCGAAAA 954
Oy 352 PheSerArgSerAlaGluThrAlaAlaAsnLeuGlnAspProAsnGlnIleuTYrAsnGlu 371
Db 955 TTCAGCAGAGATGACAGACTGCTGCCAACTGCAGAGACCCCAACGACTCTACAAATGAG 1014
Oy 372 LeuAsnLeuGlyArgArgGluGlyTYrAspValIleuGlnLYsLYsArgAlaArgAspPro 391
Db 1015 CTCAATCTAGGCGCAAGAGAGAAATACACTCTTGGAGAGAGAGCGGCTCGGAGATCCA 1074
Oy 392 GluMetGlyGlyLYsGlnIYArgArgArgAsnProGlnGlnIYValTYrAsnAlaLeu 411
Db 1075 GAGATGGAGAGCAACAGCAGAGAGAGAGAACCCCGAGAGGCAATATCAATGACACTG 1134
Oy 412 GlnLYsAspLYsMetAlaGluAlaTYrSerGluIleGlyThrLYsGlyGluArgArg 431
Db 1135 CAGAAAGACCAATGACAGAGCCTACAGATGAGATCGCACAAAGGCGAGAGCGGAGAGA 1194
Oy 432 GlyLYsGlyHisAspGlyLYsGlnIYLeuSerThrAlaThiLYsAspThrTYrAsp 451
Db 1195 GGCAAGGGGACGATGGCTTTTACAGGGTCTCAGCACTGCCACCAAGAGACCTATGAT 1254
Oy 452 AlaLeuHisMetGlnThrLeuAlaProArg 461

```

```

Db 1255 GCCCTGATATGACAGCCCTGGCCCTCGC 1284
RESULT 4
LOCUS A22539 2012 bp DNA linear PAT 15-NOV-1994
DEFINITION M.musculus/P. aeruginosa Fv(FRP5)-ETA fusion protein.
ACCESSION A22539
VERSION A22539.1 GI:641549
KEYWORDS
SOURCE
ORGANISM
other sequences; artificial sequences.
REFERENCE
1 (bases 1 to 2012)
Wels, W.S., Hynes, N.E., Harwerth, I.M., Groner, B., Hardman, N. and
AUTHORS Zwickl, M.
TITLE Recombinant antibodies specific for a growth factor receptor
JOURNAL Patent: EP 0502812-A 10 09-SRP-1992;
CIBA-GEIGY AG
FEATURES
source
location/Qualifiers
1..2012
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
1..1911
/product="Fv heavy and light chain variable domain and
exotoxin A fusion protein"
ORIGIN
Alignment Scores:
Pred. No.: 7,796-93 Length: 2012
Score: 1315.50 Matches: 280
Percent Similarity: 68.34% Conservative: 20
Best Local Similarity: 63.78% Mismatches: 45
Query Match: 53.89% Indels: 95
DB: Gaps: 10
US-09-596-774-6 (1-461) x A22539 (1-2012)
Oy 15 LYsValProLYsGlnIleGlnLeuValGlnSerGlyProGlnLeuLYsProGlyGlu 34
Db 85 AAGCTAGCTTCTCAGGATCAACACTGCAGCGAGCTGAGCTGAACATGAAGACCTGGAGAG 144
Oy 35 ThrValIleIleSerCysLYsAlaSerGlyTYrProPheThrPheThrTYrGlyMetAsnTrp 54
Db 145 ACGTCAAGATCTCTCGCAAGCGCTCTGAGTATCTTCACAAACTATGGAATGAATCTGG 204
Oy 55 ValLYsGlnAlaProGlyGlnGlyLeuLYsTrpMetGlyTYrIleAsnThrSerThrGly 74
Db 205 GTGAGAGAGCTCCAGAGACAGGCTTTAAAGTGATGGGCTGATTAACACCTCCACTGGA 264
Oy 75 GluSerThrPheAlaAspAspPheLYsGlyArgPheAspPheSerLeuGluThrSerAla 94
Db 265 GAGTCAACATTGCTGATGACTTCAAGGACAGGCTTGACTTCTCTTGGAAACCTTGCC 324
Oy 95 AspThrAlaTYrLeuGlnIleAsnAsnLeuLYsSerGluAspMetAlaThrTYrPheCys 114
Db 325 AACACTGCTTATTGCAAGATCAACAACTCAAAAGTGAAGACATGGCTACATATTCTGT 384
Oy 115 AlaArgTYrGluValTYrHisGlyTYrValProTYrTYrGlnGlnIYThrThrValThr 134
Db 385 GCAAGATGGAGGATTACACCGGCTACGTTCTTACTTGGGGCCAAAGGACACGAGTCAAC 444
Oy 135 ValSerSerGlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlySeraspIle 154
Db 445 GTTCTCTCGGCGGTGGCGGCTTGTGTGGCGGTTCGCGGCGGTGGCGGCTTGTGACATC 504
Oy 155 GlnLeuThrGlnSerHisLYsPhelSerThrSerValGlyAspArgValSerIleThr 174
Db 505 CAGCTACCAAGTCTCAAAATTTCTGTCACTTCAATGAGAGACAGGATCGACATCAC 564
Oy 175 CysLYsAlaSerGlnAspValTYrAsnAlaValAlaTYrTYrGlnIYLYsProGlyGln 194

```

```

Db      565 TCCAAAGCCAGTCAGGATGTGTATATGCTGTTCCTGATCAACAGAAACCAGACAA 624
Qy      |||
Db      195 SerProlybLeuLeuIleTyrSerIaSerArgTyrThrGlyValProSerArgPhe 214
Qy      |||
Db      625 TCTCCTAAACTTCTGATTTACTCGGCATCTCCCGGTACACAGATCCCTTCTGCTTC 684
Qy      |||
Qy      215 ThrGlySerGlySerGlyProAspPheThrPheThrIleSerSerValGlnaGluAsp 234
Qy      |||
Db      685 ACTGCAGATGGCTCTGGCGCGGATTTCACTTCCACATCAGCAGTGTGACAGGCTGAAGAC 744
Qy      |||
Qy      235 LeuAlaValIlyrPheCysGlnGlnIlePheArgThrProPheThrPheGlySerGlyThr 254
Db      CTGCAGATTATTTCTGTGACAAACATTTTCCTACTCCATTCACGTTCCGCTCGGGAGCA 804
Qy      |||
Qy      255 LysLeuGlnIleLysAlaLeuGln-----IleSerAsnSerValMetYrPheSer 271
Db      |||
Qy      805 AAATTGGAGATCAAAAGCTCTAGAGCGCGGACCTGGCGCGCTGACCGCGACAGAGGC 864
Qy      |||
Qy      272 SerValValProValLeuGlnLysValAsnSerThrThrLys----- 286
Db      |||
Qy      865 TGCCACCTGCCGCTG-----GAGACTTTCACCCGTCATCGCCAGCCGCGC 909
Qy      |||
Qy      287 -----ProValLeuArgThrProSerProValHis 296
Db      |||
Qy      910 GGCTGGAAACAACGTGACAGTCGCGCTATCCGCTGACAGCGCTGGT-CGCTCTCA--- 965
Qy      |||
Qy      297 ProThrGlyThrSer-----GlnProGlnArgProGlnuAspCysArgProArgGlySer 314
Db      |||
Qy      966 CCTGGCGGGCGACCTGTCAATGACACAGGTGACGAGTGATCCGCAAGCCCTGCGCCAG 1025
Qy      |||
Qy      315 ValIlybGlyThrGlyLeuAspPheLeuGlnuAspProLysLeuCysThrLeuLeuAspGly 334
Db      |||
Qy      1026 CCCCCGACGGCGCGCGACCTGGCGGACGATCCG----- 1061
Qy      |||
Qy      335 IleuPheIleTyrGlyValIleIleThrAlaLeuTyrLeuArgAlaLysPheSerArg 354
Db      |||
Qy      1062 -----CGA 1064
Qy      |||
Qy      355 SerIaGlnThrAlaAlaAsnLeuGlnAspProAsnGlnLeuTyrAsnGlnuLeuAsnLeu 374
Db      |||
Qy      1065 GCAGCGGAGACGCGCCGCTGCGCCCTGACCT----- 1097
Qy      |||
Qy      375 GlyArgTArgGlnuGlyTyrAspValLeuGlnuLysArgAlaArgAspProGluMetGly 394
Db      |||
Qy      1098 GCGCCGCCCGCCAGAG-----CGAGCCCTTCCTCG-----GCA 1130
Qy      |||
Qy      395 GlyLysGlnIleArgArgArgAsnProGlnGlnuGlyValTyrAsnAlaLeuGlnuLysAsp 414
Db      |||
Qy      1131 GGGACACCGGCAACGACGA----- 1148
Qy      |||
Qy      415 LysMetAlaGlnAlaTyrSerGlnuIleGlyThrLysGlyGlnuArgArgGlyLys 433
Db      |||
Qy      1149 -----GGCCGCGCGCGGCAACGCGGACGCTGTGTAG 1178
Qy      |||

RESULT 5
128526
LOCUS      128526          2012 bp          DNA          linear          PAT 06-FEB-1997
DEFINITION Sequence 13 from patent US 5571894.
ACCESSION 128526
VERSION    128526.1 GI:1819302
KEYWORDS
SOURCE     Unknown.
ORGANISM   Unknown.
REFERENCE 1 (bases 1 to 2012)
AUTHORS   Wells,W.S., Hyman,N.E., Harweth,I.-M., Groner,B., Hardman,N. and
           Zwickl,M.
TITLE     Recombinant antibodies specific for a growth factor receptor
JOURNAL   Patent: US 5571894-A 13 05-NOV-1996;
FEATURES   Location/Qualifiers
           1..2012
           /organism="Unknown"
           /mol_type="unassigned DNA"

```

```

ORIGIN
Alignment Scores:
Pred. No.:      7,796-93
Score:          1315.50
Percent Similarity: 68.34%
Best Local Similarity: 63.78%
Query Match:    53.89%
DB:              6
Gaps:            10

US-09-596-774-6 (1-461) x 128526 (1-2012)
Qy      15 LysValProLysGlnIleGlnLeuValGlnSerGlyProGlnuLeuLysAspProGlyGln 34
Db      |||
Qy      85 AAGTACAGTCTTCAGGTACAGTCACTGACAGTCTGACCTGAAGTCAAGAGCTGAGAG 144
Db      |||
Qy      35 ThrValLysIleSerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTrp 54
Db      |||
Qy      145 ACAGTCAAGATCTCTGCAAGGCTCTGGGTATCTCTTCCACAACTATGAAATGAACTGG 204
Qy      |||
Qy      55 ValIlybGlnAlaProGlyGlnGlyLeuLysTrpMetGlyTrpIleAsnThrSerThrGly 74
Db      |||
Qy      205 GTGAAGCAGGCTCCAGACAGGCTTTAAAGTGAAGGCTGAGATTAACACCTCCAGTGA 264
Qy      |||
Qy      75 GluSerThrPheAlaAspAspPheLysGlyArgPheAspPheSerLeuGlnuThrSerAla 94
Db      |||
Qy      265 GAGTCAACATTTCTGATGACCTTCAAGGAGCGTTTGACTTCTCTTGGAAACCTCTGCC 324
Qy      |||
Qy      95 AsnThrAlaTyrLeuGlnIleAsnAsnLeuLysSerGlnuAspMetAlaThrTyrPheCys 114
Db      |||
Qy      325 AACACTGCTATTGTCAGATCAACAACTCAAAAGTGAAGCATGCTACATATTTCTGT 384
Qy      |||
Qy      115 AlaArgTrpGlnuValTyrHisGlyTyrValProTyrTrpGlyGlnGlyThrThrValThr 134
Db      |||
Qy      385 GCAAATGAGAGGTTTACACGCTTACCTTACTGAGGCGCAAGGACACAGGTCAC 444
Qy      |||
Qy      135 ValSerSerGlyGlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlySerAspIle 154
Db      |||
Qy      445 GTTTCCTCTGGCGGCGGCGGCTTCTGTGTGGCGTGCCTCGGCGGCGGCTTCTGACATC 504
Qy      |||
Qy      155 GlnLeuThrGlnSerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThr 174
Db      |||
Qy      505 CAGCTGACCCAGCTCTCAAAATTCCTGTCACCTAGTAGAGAGACAGGCTCAGCATCAC 564
Qy      |||
Qy      175 CysLysAlaSerGlnAspValTyrAsnAlaValAlaTyrPyrGlnGlnuLysProGlyGln 194
Db      |||
Qy      565 TCCAAAGCCAGTCAGAGTGTATATGCTGTGCTGTGATCAACAGAAACCAGACAA 624
Qy      |||
Qy      195 SerProlybLeuLeuIleTyrSerIaSerArgTyrThrGlyValProSerArgPhe 214
Db      |||
Qy      625 TCTCCTAAACTTCTGATTTACTCGGCATCTCCGATCACCTGGAGTCCCTTCTGCTTC 684
Qy      |||
Qy      215 ThrGlySerGlySerGlyProAspPheThrPheThrIleSerSerValGlnaGluAsp 234
Db      |||
Qy      685 ACTGCAGTGGCTCTGGCGCGGATTTCACTTCCACATCAGCAGTGTGACAGGCTGAAGAC 744
Qy      |||
Qy      235 LeuAlaValIlyrPheCysGlnGlnIlePheArgThrProPheThrPheGlySerGlyThr 254
Db      |||
Qy      745 CTGCAGATTATTTCTGTGACAAACATTTTCCTACTCCATTCACGTTCCGCTCGGGAGCA 804
Qy      |||
Qy      255 LysLeuGlnIleLysAlaLeuGln-----IleSerAsnSerValMetYrPheSer 271
Db      |||
Qy      805 AAATTGGAGATCAAAAGCTCTAGAGCGCGGACCTGGCGCGCTGACCGCGACAGAGGC 864
Qy      |||
Qy      272 SerValValProValLeuGlnLysValAsnSerThrThrLys----- 286
Db      |||
Qy      865 TGCCACCTGCCGCTG-----GAGACTTTCACCCGTCATCGCCAGCGCGGC 909
Qy      |||
Qy      287 -----ProValLeuArgThrProSerProValHis 296
Db      |||
Qy      910 GGCTGGAAACAACGTGACAGTCGCGCTATCCGCTGACGAGCGCTGGT-CGCTCTCTA--- 965
Qy      |||
Qy      297 ProThrGlyThrSer-----GlnProGlnArgProGlnuAspCysArgProArgGlySer 314

```

Db 966 CCTGGCGCGCGCGCTCATGCAACGAGTCCGACGAGTATCCGCAACGCCCTGGCCAG 1025
QY 315 ValLysGlyThrGlyLeuAspPheLeuGluAspProLysLeuGlyLeuLeuAspGly 334
Db 1026 CCCCAGCGCGCGCGCGCTGGCGCAACGATCCG----- 1061
QY 335 IleLeuPheIleTyrGlyValIleIleThrAlaLeuTyrLeuArgAlaLysPheSerArg 354
Db 1062 -----CGA 1064
QY 355 SerAlaGluThrAlaAlaAsnLeuGluAspProAsnGlnLeuTyrAsnGlnLeuAsnLeu 374
Db 1065 GGAGCGCGAGCAGCGCCCTGCGCCCTGACCT----- 1097
QY 375 GlyArgArgGluGluTyrAspValLeuGluLysLysArgAlaArgAspProGluMetGly 394
Db 1098 GCGCGCGCGCGAGC-----CGAGCGCTTCGTCG-----GCA 1130
QY 395 GlyLysGlnGlnArgArgAspProGlnGluGlyValTyrAsnAlaLeuGlnLysAsp 414
Db 1131 GGGCAGCGCGCAACGACGA----- 1148
QY 415 LysMetAlaGluAlaTyrSerGluIleGlyThrLysGlyGluArgArgGlyLys 433
Db 1149 -----GGCGCGCGCGCGCAACGCCGACGTGTGAG 1178
RESULT 6
A57337 2070 bp DNA linear PAT 03-MAR-1998
LOCUS Sequence 7 from Patent EP0739984.
DEFINITION A57337
ACCESSION A57337.1 GI:3713216
VERSION A57337.1
KEYWORDS
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE 1. Weis, W.D., Schmidt, M. and Groner, B.D.
AUTHORS Bivalent polypeptides containing at least two domains
TITLE Patent: EP 0739984-A 7 30-OCT-1996;
JOURNAL SAN TUMORFORSCHUNGS GMBH (DE)
FEATURES
source Location/Qualifiers
1..2070
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/clone="PSM202-5"
1..1956
/note="unnamed protein product"
/codon_start=1
/transl_table=1
/protein_id="CAA03441.1"
/db_xref="GI:3713217"
/translation="MKRTATATIAVALAGFATVAQADYKDDDKLHHHHKLOVOLOO
SGPELKRGETVTKISCKASGPTNTYGNWVKAPGGGLKMMWINTSSESTPADP
KGRDFSLSTSAANTAYLIQINILKSDMAFYFCARMEYHGVPTWGGITVTVS
GSGGSGGSGGSDIQLTQSHKFLSTSDVRSITCKASQDVYNAVAYOQKPGSPYL
LIYASSRYTGVPSRFPTGSGGSDPFTSTISVOAEDLAVYFCQHPPTPTFGSTGL
EIKALEHHHHHLEGGSLAALTAAHQAQCHLPLETTRHRORGRMGLECCQVPORYVA
LYLAARLSMNQDVIVRNALASPGSGDLGAIHREPEQARLAIITLAAASSERFVRG
TGNDENAGANADVSLTCPVAAGCGAPADGDLBERNPPTGAEFLDGDADVFSFR
GTQWTVVERLQAAROLEBERGVFVGHGTFLEAASIVFGVAVARQDDAIWRGY
IAGDPALAYGAQDOEDDARGLRNGALLRVYRPSLPGYRPSLTILTAPEAAGEVE
RLTGHPPLRLDAITGPEEGGRLETTILGWPLEARIVIPSAIPTDRNVGGLDPS
IPDKQAISALPDYASQPKPKPRDLK"

ORIGIN
Alignment Scores:
Pred. No.: 1.38e-92 Length: 2070
Score: 1312.50 Matches: 284
Percent Similarity: 68.59% Conservative: 13
Best Local Similarity: 65.59% Mismatches: 55

Query Match: 53.77% Indels: 83
DB: 6 Gaps: 8
US-09-596-774-6 (1-461) x A57337 (1-2070)
QY 19 GlnIleGlnLeuValGlnSerGlyProGluLeuLysLysProGlyGluThrValLysIle 38
Db 115 CAGGTACAACTGACGACAGTCTGACCTGAACCTGAAGAAAGCTGAGACAGTCAAGTC 174
QY 39 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValLysGlnAla 58
Db 175 TCTGCAAGGCGCTCGGGTATCTTCAAAACATAGAAATGAACCTGGTGAAGCAGGCT 234
QY 59 ProGlyGlnGlyLeuLysTyrMetGlyTyrTrpLysSerThrArgLysGlyLysThrPhe 78
Db 235 CCAGACAGGCTTTAAAGTGGATGGGCTGGATTAACTTCACTCCACTGAGAGTCAACATTT 294
QY 79 AlaAspAspPheLysGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 98
Db 295 GCTGATGACTTCAGAGGACGGTTTGACTTCTTTGAAACCTTGGCAACACTGCTAT 354
QY 99 LeuGlnIleAsnAsnLeuLysSerGluAspMetAlaThrTyrPheCysAlaArgTrpGlu 118
Db 355 TTGCAGATCAACAACTCAAAAGTGAAGACATGGCTACATATTCTGTGCAAGATGGAG 414
QY 119 ValTyrHisGlyTyrValProTyrTrpGlyGlnGlyThrThrValThrValSerSerGly 138
Db 415 GTTACACAGCGCTACGTTCTTCTGAGGACAGGACACCGGTACCGTTCTCTGGC 474
QY 139 GlyLysGlySerGlyGlyGlySerGlyGlyGlySerAlaIleGlnLeuThrGln 158
Db 475 GGTGGCGGTTCTGGTGGCGGTGGCTCCGCGGCGGCTTCTGACATCCAGTCAACCCAG 534
QY 159 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSer 178
Db 535 TCTCAAAATTCCTGTCCTCACTCAGTAGAGACAGGCTCAGATCACTGCAAGGCCAGT 594
QY 179 GlnAspValTyrAsnAlaValAlaTyrGlyGlnLysProGlyLysSerProLysLeu 198
Db 595 CAGAGTGTATATATGCTGTGCTGCTGTATCAACAGAACAGACATCTCTTAACCT 654
QY 199 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 218
Db 655 CTGATTTACTCGGCATCTCTCCGCTACCTGAGTGCCTTCTGCTTCACTGGCAGTGGC 714
QY 219 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 238
Db 715 TCTGGCGCGGATTTCACTTTCACCATCAGCATGTGTGACGCTGAAGACCTGGCAGTTTAT 774
QY 239 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyThrLysLeuGluIle 258
Db 775 TTCTGTCAGCAACATTTTGTACTTCATTCACCTTGGCTCGGGGACAAATTTGGAGATC 834
QY 259 LysAlaLeuGluLysSerAsnSerValMetTyrPheSerSerValValProValLeuGln 278
Db 835 AAAGCTCTAGAGCA----- 848
QY 279 LysValAsnSerThrThrThrLysProValLeuArgThrProSerProValHisProThr 298
Db 849 -----CAATCATCAACCATCACTAG 869
QY 299 GlyThrSerGlnProGlnArgProGluAspCysArgProArgGlySerValLysGlyThr 318
Db 870 GGG---CGGACGCTGGCGCGCGTGC-----CGGCACACAGGCTCGCACTGGCGCT 920
QY 319 GlyLeuAspPheLeuGlnAspProLysLysCysTyrLeuLeuAspGlyTyrLeuPheIle 338
Db 921 GGAGACTTTCACCGCTCATCGCCAGCGCGCGCTGGGAACAATCGACAGTGGCGCTA 980
QY 339 TyrGlyValIleIleThrAlaLeuTyrLeuArgAlaLysPheSer----- 353
Db 981 TCCGATGCAAGCGGCTGTGCTGCTTCTACTGGCGGCGGACATGTATGAAACAGAGTCA 1040

QY 354 -----ArgSerAlaGluThrAla-----AlaAs 361
|||
DB 1041 CCAGGTGATCCCAACGCCCTTGGCCAGCCCGGCGGCGGAGCCTGGCGGAAGCGAT 1100
|||
QY 361 nLeuGlnAapProaenGlnLeuTyraenGlnLeuAenGlnIyaArgGlnGlyTras 381
|||
DB 1101 CCGCAGACGCGCGGAGCGCCGCTCTGGCCCTGACCTT-GGCGCGCGCGGAGAG----- 1154
|||
QY 381 pValLeuGlnIySerAaArgAaProGlnuMetGlyGlySerGlnGlnAaArgAaR 401
|||
DB 1155 -----CGAGCGCTTCTGCTCG-----GCAGGCGACCGCGCAAGCAAG- 1190
|||
QY 401 GAaapProGlnGlnGlyValTyraAaAlaLeuGlnIySerAaPheMetAlaGluAlaTyse 421
|||
DB 1190 ----- 1190
|||
QY 421 rGlnIleGlyThrIySerGlyGlnAaArgAaGlyIySer 433
|||
DB 1191 -----GGCGCGCGCGCGCAACGCGGAGCTGTGTAG 1220
|||
RESULT 7
LOCUS A57353 2730 bp DNA linear PAT 03-MAR-1998
DEFINITION Sequence 23 from Patent EP0739984.
ACCESSION A57353
VERSION A57353.1 GI:3713232
KEYWORDS
SOURCE
ORGANISM
SYNTHETIC CONSTRUCT
SYNTHETIC CONSTRUCT
OTHER SEQUENCES: artificial sequences.
REFERENCE
AUTHORS 1
TITLE Weis W.D., Schmidt M. and Groner B.D.
JOURNAL Bivalent polypeptides containing at least two domains
SAN TUDORFORSCHUNGS GMBH (DE)
LOCATION/Qualifiers
1. 2730
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/clone="PMS246-5-5 KDEL"
1. 2700
/note="unnamed protein product"
/codon_start=1
/transl_table=1
/protein_id="CAA03449.1"
/db_xref="GI:3713233"
/translation="MKRTAIAIAVALAGFATVAQADYKDDDKLHHHHHKLQVQLQ
SGPELKKRGETVTKISCKASGYPTNYGMWVQAQGGGLKMMGINTSTGSPADDF
KGRDPSLETSANPAVLIQINLKSEDMATYCAPEWYHGVVPGVGGCTTYVSGGG
GSGGSGSGGSDIOLTOSHKPLSTSVGDRVSTTKASQDVYNAVAMTQKPGSPKL
LIYASRSYTVSPSRFTGSGGSDPTFTISSVQAEDLAVPCQOHFRTPFTFGSTKL
EIKALEHHHHHLEGGSLAALTALHQAQCHLPLETFRHROPFGMEQLSCCGPVQGLA
LYLAARLSMNQDQVIRNALASPGSGDLGEAIRQPPQARLALTLAAESERFVQG
TGNDEAGANADVSLTGPVAAEGCAGSDGLDENYRGAFLDGLDVSFSTR
GTQWVYERLLQARHOLEBERGVPCYHGTLEAAGSYRFGVARSDDLDLIRGFI
IAGRPALAYGTAQDEDPARKIRNGALLRIVPRSSLPGRTRISLTILAAEPAGEVE
RLIGHPLRLDAITGPREEGRLETLIGWPLAERTVVIISAIPDPNNVGGDLPSS
IPDKEOALISALPDVASQPKPREASQVLOQSGPELKKRPGSTVTKISCKASGYPTNY
GMWVQAQGGGLKMMGINTSTGSPADDFKGRDPSLETSANPAVLIQINLKSEDM
ATYCAPEWYHGVVPGVGGCTTYVSGGGSGGSDIOLTOSHKPLSTSVGDRVSTTK
ASQDVYNAVAMTQKPGSPKL
VGDVSTTKASQDVYNAVAMTQKPGSPKL
FTISSVQAEDLAVPCQOHFRTPFTFGSTKL
EIKALEKDEL"

US-09-596-774-6 (1-461) x A57353 (1-2730)
QY 19 GlnIleGlnLeuValGlnSerGlyProGlnLeuIySerProGlyGluThrValIyIle 38
|||
DB 115 CAGGTACACAGTCAGACAGCTGAGCTTGAACCTGAAGAGCCTGGAGAGACAGTCAAGATC 174
|||
QY 39 SerCySerAaIaSerGlyTyraProPheThrAenTyrgIyMetAenTyraIySerGlnA 58
|||
DB 175 TCCTGCAAGGCTCTGGGTATCTCTTCCAAACATAGTAAGAACCTGGGTGAAGAGGCT 234
|||
QY 59 ProGlyGlnIyLeuIySerTyraMetGlyTyraIleAenTyraSerThrGlyGluSerThrPhe 78
|||
DB 235 CCAGGACAGGCTTAAAGTGAAGGCTGAGATTAAACCTCACCTGAGAGATCAACATTT 294
|||
QY 79 AlaAaAapPheIySerGlyAaArgAaPheSerLeuGlnuThrSerAlaAenThrAlaTy 98
|||
DB 295 GCTGATGCTTCAAGGAGCGGTGACTTCTTTGAAACCTGCGCAACAGCTGCTAT 354
|||
QY 99 LeuGlnIleAaAenLeuIySerGlyAaAaPheAlaThrTyraPheCyAaIaAaGTTPGlu 118
|||
DB 355 TTGCAGATCAACACCTCTCAAAAGTGAACATGGCTACATTTCTGTGCAAGATGGAG 414
|||
QY 119 ValTyraIySerGlyTyraValProTyraTyraGlyGlnGlyThrValIySerSerGly 138
|||
DB 415 GTTTAACACAGGCTACGTTCTTACTGGGGCCAAAGGACCAAGCTCACCGTTCTCTGGC 474
|||
QY 139 GlyGlyIySerGlyGlyGlyGlySerGlyGlyGlyIySerAaPheIleGlnLeuThrGln 158
|||
DB 475 GGTGGCGGTGTGGGGGGGGTGGCTCCGGCGGTGGCTTGCATATCAGCTGACCGAG 534
|||
QY 159 SerIyIyPheLeuSerThrSerValGlyAaAaPheValSerIleThrCySerAaIaSer 178
|||
DB 535 TCTCAAAATCTCTGTCCATCTTCACTGAGAGACAGAGTCAACATCACTGCAAGCCAGT 594
|||
QY 179 GlnAaAaIyTyraAaIaValAlaTyraGlyGlnGlnIySerProGlyGlnIySerProIySer 198
|||
DB 595 CAGGATGTGATTAATGCTGTGCTGATCAACAGAAACAGGACCAATCTCTTAACCT 654
|||
QY 199 LeuIleTyraSerAlaSerSerAaTyraThrGlyValProSerAaArgPheThrGlySerGly 218
|||
DB 655 CTGATTAATCTGGACATCTCCCGGACCTGAGTCCCTTCTGCTTACGTGGCGTGGC 714
|||
QY 219 SerGlyProAaPheThrPheThrIleSerSerValGlnAaGluAaPheAlaValTy 238
|||
DB 715 TCTGGCGCGGATTTCACTTCACTTCACTGAGAGTGGAGGCTGAAGACTGGCACTTAT 774
|||
QY 239 PheCySerGlnGlnIyAaPheArgThrProPheThrPheGlySerGlyTyraIyLeuGlnIle 258
|||
DB 775 TTCTGTGCAAGCAATTTTGTATCTCAATTCAGTTCCGCTCGGGGACAAATTTGGAGATC 834
|||
QY 259 LysAlaLeuGlnIleSerAaSerValMetTyraPheSerSerValIyProValLeuGln 278
|||
DB 835 AAAGCTCAGAGCA----- 848
|||
QY 279 LysValAaSerThrThrThrIySerProValLeuAaTyraThrProSerProValIyProThr 298
|||
DB 849 -----CATCACTCACTCACTCACTGA 869
|||
QY 299 GlyTyraSerGlnProGlnAaArgProGluAaPheCyAaArgProGlySerValIySerGlyThr 318
|||
DB 870 GGG---CGGAGCTGGCGCGGCTGAC-----CGGCAACAGGCTGCCACCTGCCCT 920
|||
QY 319 GlyLeuAaPheLeuGlnAaPheProIyLeuCySerTyraLeuAaPheGlyIyLeuPheIle 338
|||
DB 921 GGAAGCTTTCACCCGCTCATCCGAGCGCGGCTGGAGAACATCGAGAGCTGGCGCTA 980
|||
QY 339 TyrGlyValIle-IleThrAlaLeuTyraIySerAaIyAaPheSer----- 353
|||
DB 981 TCCGTGACAGGCGCTGCTGCTTCTTACTGCGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1040
|||
QY 354 -----ArgSerAlaGluThrAla-----AlaAs 361
|||
DB 1041 CCAGGTGATCCCAACGCCCTTGGCCAGCCCGGCGGCGGAGCCTGGCGGAAGCGAT 1100
|||

QY 361 nLeuGlnAProAsnGlnLeuTyrAsnGlnLeuAsnLeuGlyAArgGlnGluTyrAs 381
 Db 1101 CCGCGAGCGCCGAGAGCCCGCTGGCCCTGACCTT-GGCGCGCGCGAGAG----- 1154
 QY 381 pValleuGlnLysArgAlaArgAspProGlnuMetGlyLysGlnGlnArgArgAr 401
 Db 1155 -----CGAGCGCTGCTCCG-----GCAGGGGCAACCGGCAACGAGCA-- 1190
 QY 401 gAsnProGlnGlnGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaGlnAlaTyrSe 421
 Db 1190 ----- 1190
 QY 421 rGluIleGlyThrLysGlyLysArgArgGlyLys 433
 Db 1191 -----GGCGCGCGCGCGCAACCGGAGCTGTGTAG 1220

RESULT 8

XXU64991

LOCUS

894 bp DNA linear SYN 11-FEB-2002

DEFINITION Synthetic construct anti-human erbB-2 receptor bifunctional

single-chain antibody-phosphatase fusion protein, partial cds.

ACCESSION U64994.1 GI:1545865

KEYWORDS

SEGMENT

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

FEATURES

SOURCE

CDS

ORIGIN

ALIGNMENT SCORES:

Pred. No.:

Score:

Percent Similarity:

Best Local Similarity:

Query Match:

DB:

US-09-596-774-6 (1-461)

X XXU64991 (1-894)

1.08e-92

1308.00

94.32%

93.18%

53.58%

12

894

246

3

3

12

1

894

246

3

3

12

1

894

246

3

3

12

1

894

QY 15 LysValProLysGlnIleGlnLeuValGlnSerGlyProGlnuLeuLysLysProGlyLys 34
 Db 85 AAGCTAGCTTCTCAGATACCACTGACAGAGCTGAGCCTGAGACGAGAACCTGAGAG 144

QY 35 ThrValLysIleSerCysValLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyr 54
 Db 145 ACACTGACATCTCCTGAGAGGCTCTG3ATATCTTCCACAACTATGAGATGACCTG 204

QY 55 ValLysGlnAlaProGlyGlnGlyLysLysLysTyrPheGlyTyrPheThrSerThrGly 74
 Db 205 GTGAGGACAGGCTCAGAGCAGAGGTTAAAGTGAAGTGGCTGAGATTAAACCTCAGC 264

QY 75 GluSerThrPheAlaAspAspPheLysGlyLysPheAspPheSerLeuGlnThrSerAla 94
 Db 265 GAGTCACATTTGCTGAGTACTTCAAGAGACGTTTGACTTCTTGGAAACCTGCTGCC 324

QY 95 AsnThrAlaTyrLeuGlnIleAsnAsnLeuLysSerGlnuLysPheMetAlaThrTyrPheCys 114
 Db 325 AACACTGCTATTTCAGATACCACTGACCAAAAGTGAAGACATGCTACATATTCTCT 384

QY 115 AlaArgTyrGlnValTyrLysGlyTyrValProTyrTyrGlnGlnGlyThrValThr 134
 Db 385 GCAAGATGGAGGTTTACACGCGCTACGCTTACTGAGGCGCAAGGAGCACCGTCA 444

QY 135 ValSerSerGlyLysGlyLysSerGlyLysGlyLysSerGlyLysGlyLysSerAspIle 154
 Db 445 GTTTCCTCTGCGGAGGCGGTTCTGTGCGAGTGGCTCCGCGGCGGCTTCTGACATC 504

QY 155 GlnLeuThrGlnSerHISLysPheLeuSerThrSerValGlyAspArgValSerIleThr 174
 Db 505 CAGCTGACCCAGCTCTCAAAATTCCTCTGCTCACTGAGAGACAGGTCAGATCACC 564

QY 175 CysLysAlaSerGlnAspValTyrAsnAlaValAlaTyrGlnGlnLysProGlyGln 194
 Db 565 TGCAAGGCGAGTCAAGATGTGTAAATGCTGTGCGCTGATCAACAGAAACAGAGCA 624

QY 195 SerProLysLeuLeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPhe 214
 Db 625 TCTCTTAACTTCTGATTTACTGGCATCTCCCGGATCTGAGATGCCCTTCTGCTTC 684

QY 215 ThrLysSerGlySerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAsp 234
 Db 685 ACTGCGAGTGCCTCGGCGGAGATTCACTTCACTGACATGAGAGTGCAGGCTGAGAG 744

QY 235 LeuAlaValTyrPheCysGlnGlnIlePheAspArgThrProPheThrPheGlySerGlyThr 254
 Db 745 CTGGCACTTATTTCTGTGACAACTTTCTGATCTCATTCACGTTTGGCTCGGGGACA 804

QY 255 LysLeuGlnLysAlaLeuGlnIleSerAsnSerValMetTyrPheSerSerValVal 274
 Db 805 AAATTGAGATCAAGCTTAGAG----- 828

QY 275 ProValLeuGln 278
 Db 829 CCTGTTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

QY 829 CCGTCTCTGAAA 840

source
1. 1862
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
CDS
1. 1854
/note="unnamed protein product"
/codon_start=1
/protein_id="CA03250.1"
/db_xref="GI:2303795"
/translation="MDYKDDDDKKLHHHHHKLSSISBOACDICKLKSKSEKPK
AKLKNNECKRSPKTKRSLPLRAHLTEVESLERLEQLFLIPREDIDMLKQDLSL
QDIALKALGLFQDNDVNDKDAVTDRLAETDMLPLRLRHRISATSSSESSKQROL
TVSSSLAVGSSLSICINLMDVDIRDKTKIKLSKEHGPDKNMSSESPKTKYSEKAKO
LLEEFHOTALEHEPELSELTGTNPVPAAGAYAAVAVAOVIDSEADNLEKTTAA
LSILPDISVWGILADGAVHNHNEEIVAOISILSSIMVAOAIPLVGELVDIGPAAYNV
ESILNFOVHNHSYRPRAYSPGVSDIDKLOVLOQSGELKPKPGTIVKISCASTP
TNYGNMVKQAPGQGLKMMGMTINTSTGSTFDDPKGRDFSLTSANTAYLIQNNLK
SEDMATYFCARWEVYHGVYVYWGQGTITVTSVSSGSGSGSGSGSDIQLTQSHKFL
STVGDRAVITCKASQADYNAVAVAOQKPGSGKLLIYSASRYGVSFRFGSGSGSP
DPTFTISSVQAEADLAVYFCQGHFRTPFGSTKLEIKALEDLSESRFSA"

ORIGIN

Alignment Scores:
Pred. No.: 2,71e-92 Length: 1862
Score: 1308.00 Matches: 245
Percent Similarity: 97.64% Conservative: 3
Best Local Similarity: 96.46% Mismatches: 6
Query Match: 53.58% Indels: 0
DB: 6 Gaps: 0
US-09-596-774-6 (1-461) x AS0095 (1-1862)

Qy 19 GlnIleGlnLeuValGlnSerGlyProGluLeuValSerProGlyGluThrValIle 38
Db 1090 CAGGTACACCTGACAGCTGAGCTGAACTGAAAGACCTGAGAGACACTCAAGATC 1149
Qy 39 SerCysValAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValIleGlnAla 58
Db 1150 TCTGTCAGAGGCTCTGGGATCTTTCACAACTATGAAATGAACTGGGTGAAGCAGCT 1209
Qy 59 ProGlyGlnGlyLeuValTyrMetGlyTyrIleAsnThrSerThrGlyIleSerThrPhe 78
Db 1210 CCAGGACAGGGTTTAAAGTGAAGTGGGCTGATTAACCTCCACCTGAGAGATCAATTT 1269
Qy 79 AlaAspAspPheLeuValGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 98
Db 1270 GCTATGACTTCAAGGAGAGGTTTGAATCTTCTTGGAAACCTCGCCAACTGCTCAT 1329
Qy 99 LeuGlnIleAsnAsnLeuValSerGlyLysAspMetAlaThrTyrPheCysAlaArgTyrGlu 118
Db 1330 TTGCAGATCAACCACTCAAAAGTGAAGATGCTACATATTTCTGTCGAAGATGGAG 1389
Qy 119 ValTyrHisGlyTyrValProTyrTyrGlyGlnGlyThrValThrValSerSerGly 138
Db 1390 GTTTCACACGGCTAGCTTCTTACTGAGGACCAAGGAGACCAAGCTTCTCTGGC 1449
Qy 139 GlyIleGlySerGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThrGln 158
Db 1450 GGTGGCGGTTCTGGGCGGTGCTCCGGCGGTGCGGTTCTGACATCCAGCTGACCCAG 1509
Qy 159 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysAlaValAspSer 178
Db 1510 TCTCAAAATCTCTTCCACTTCAAGAGACAGGCTACGATCACTGCAAGGCCAGT 1569
Qy 179 GlnAspValTyrAsnAlaValAlaTyrTyrGlnGlnLysProGlyGlnSerProLysLeu 198
Db 1570 CAGAGTGTATATATGCTGCTGCTGATCAAGAAACAGAGCAATCTCTTAACAT 1629
Qy 199 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 218
Db 1630 CTGATTTACTCGGCACTCTCCCGGATACACTGAGATCCCTTCTCGCTTCACTGGCAGTGGC 1689
Qy 219 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 238

Db 1690 TCTGGGCGGATTTCACTTCACTTCAAGCTGAGCTGAGAGCTGAGACCTGGCACTTAT 1749
Qy 239 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyThrLysLeuGluIle 258
Db 1750 TTCTGTACAGACATTTTCTGATCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTT 1809
Qy 259 LysAlaLeuGlnIleSerAsnSerValMetTyrPheSerSer 272
Db 1810 AAAGCTTACAGATCTCTCGAGTGAAGAAATTTTCAAGC 1851

RESULT 10
AR268367 1862 bp DNA linear PAT 10-APR-2003
LOCUS AR268367
DEFINITION Sequence 36 from patent US 6498233.
ACCESSION AR268367
KEYWORDS AR268367.1 GI:29698717
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
1 (bases 1 to 1862)
AUTHORS
Wels,W. and Fominaaya,J.
TITLE
Nucleic acid transfer system
JOURNAL
Patent: US 6498233-A 36 24-DEC-2002;
FEATURES
LOCATION/Qualifiers
source
1. 1862
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores:
Pred. No.: 2,71e-92 Length: 1862
Score: 1308.00 Matches: 245
Percent Similarity: 97.64% Conservative: 3
Best Local Similarity: 96.46% Mismatches: 6
Query Match: 53.58% Indels: 0
DB: 6 Gaps: 0
US-09-596-774-6 (1-461) x AR268367 (1-1862)

Qy 19 GlnIleGlnLeuValGlnSerGlyProGluLeuValSerProGlyGluThrValIle 38
Db 1090 CAGGTACACCTGACAGCTGAGCTGAACTGAAAGACCTGAGAGACACTCAAGATC 1149
Qy 39 SerCysValAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValIleGlnAla 58
Db 1150 TCTGTCAGAGGCTCTGGGATCTTTCACAACTATGAAATGAACTGGGTGAAGCAGCT 1209
Qy 59 ProGlyGlnGlyLeuValTyrMetGlyTyrIleAsnThrSerThrGlyIleSerThrPhe 78
Db 1210 CCAGGACAGGGTTTAAAGTGAAGTGGGCTGATTAACCTCCACCTGAGAGATCAATTT 1269
Qy 79 AlaAspAspPheLeuValGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 98
Db 1270 GCTATGACTTCAAGGAGAGGTTTGAATCTTCTTGGAAACCTCGCCAACTGCTCAT 1329
Qy 99 LeuGlnIleAsnAsnLeuValSerGlyLysAspMetAlaThrTyrPheCysAlaArgTyrGlu 118
Db 1330 TTGCAGATCAACCACTCAAAAGTGAAGATGCTACATATTTCTGTCGAAGATGGAG 1389
Qy 119 ValTyrHisGlyTyrValProTyrTyrGlyGlnGlyThrValThrValSerSerGly 138
Db 1390 GTTTCACACGGCTAGCTTCTTACTGAGGACCAAGGAGACCAAGCTTCTCTGGC 1449
Qy 139 GlyIleGlySerGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThrGln 158
Db 1450 GGTGGCGGTTCTGGGCGGTGCTCCGGCGGTGCGGTTCTGACATCCAGCTGACCCAG 1509
Qy 159 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysAlaValAspSer 178
Db 1510 TCTCAAAATCTCTTCCACTTCAAGAGACAGGCTACGATCACTGCAAGGCCAGT 1569

Qy	179	GlnAspValTyrAsnAlaValAlaIlePheTyrGlnGluLysProGluGlnSerProLysLeu	198
Db	1570	CAGATGTGTATANGCTTGTCCTCGGTATCAACGAAACACGACCAATTTCTTAACCT	1622
Qy	199	LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly	218
Db	1630	CTGATTTACTGGCATCTCCCGGACACTGGAGTCCCTTCCTCGCTTCACTGGCAGTGGC	1685
Qy	219	SerGlyProAspPheThrPheThrTyrLeuSerValGlnAlaGluAspLeuAlaValTyr	238
Db	1690	TCTGGGCGGATTTCACTTTCACCATCAGCAGTGTGCAGGCTGAAGACTCGGCACTTAT	1749
Qy	239	PheCysGlnGlnIleAspPheArgThrProPheThrPheGlySerGlyTyrLysLeuGluIle	258
Db	1750	TTCTGTAGCAACAATTTTCGATCTCAATTCACGTTCCGCTCGGGGACAAATTTGGAGATC	1809
Qy	259	LysAlaLeuGluIleSerAsnSerValMetTyrPheSerSer	272
Db	1810	AAAGCTTAGAGGATCTTCGAGTGAGAGAAATTTTTCAGCC	1851

RESULT	11
A50993	
LOCUS	
DEFINITION	A50993
Sequence	34 bp
Accession	From Patent WO9613599.
Accession	M50003
DNA	
linear	PAT 07-MAR-1997

VERSION	A50993.1	GI:2303792
KEYWORDS	.	
SOURCE	unidentified	
ORGANISM	unidentified	

REFERENCE
AUTHORS
TITLE
JOURNAL

1 (bases 1 to 19)
Wels, W. and Fominaya, J.
NUCLEIC ACID TRANSFER SYSTEM
Patent: WO 961359-A 34 03-MAY-1996;
Microbiology

COMMENT	WELS WINFRIED (DE)
Other publication AU 3926895 960523.	
Location/Qualifiers	
1. .1919	
source	

```

/mol_type="unassigned DNA"
/db_xref="taxon:32644"
64. .1908
CDS

```

```

/note="unnamed protein product; protein sequence is in
conflict with the conceptual translation"
/codon_start=1
/protein_id="CAA03249.1"
/db_xref="GI:2303793"

```

/translacion="DYKDDDDDKHHNNHKLSSIBOACIOCRKLKSGKEXKQYK
 CLKNOMECRYSPKTKRSP/LTRALTHEVSKLRLBUEFLIPREDIDMLKXDSIOD
 IKALLTGLFVQVNDVNDKAVTRILASVETDMP/LTRQHRISLATSSESSNKKGOROLV
 SSSILASGSLSSCINLMDVDVIRDKTKIEB/KHGP/KNOMSEPNKTVSEBKAOYL
 BEHQATLEHPELSLMDTKVTGTPVFPKGAAYAMA/VNAVQIDSETDNKLETKAALS
 ILPGISGVMSIADGAVHNHTEELVIAOSIALSSLMVAQATLVGSELNDIGRAVNPRES
 IINLFOVSMSIYRPAVSPGSDGIDDKLQVLOQSGPELKKRPELTSVISCASGSPFN
 YGNWVWQAPQAGOLKPMWPGINTGBESTFADDPKGRDFELLETSAVYAOIINLKESS
 DMATYFCARWEVNHVGYVPMWGCGTIVYSSGGGSGGGGSGGGGSDIOLQSHKPLST
 SVYDRUSITCKASODVYNAVAWQOKRGSOKPLIYASRRTYGVSPRFGSGSGPDF
 TFLISSVQABDLAVYFCQGHFRTPFTTGGSTKLEIKALEDLSSBRPSA"

	Alignment Scores:	
Pred. No.:	2.82e-92	Length: 1919
Score:	1308.00	Matches: 245
Percent Similarity:	97.64%	Conservative: 3
Best Local Similarity:	96.46%	Mismatches: 6
Query Match:	53.58%	Indels: 0
B:	6	Gaps: 0

US-09-596-774-6 (1-461) X A50993 (1-1919)

QY 19 GlnIleGlnLeuValGlnSerGlyProGlnLeuIleLysIleProGlyGluThrValLysIle 38
|||:::|||||
1147 CAGGTCACACTGCAGCAGTCTGGACCTGAACTAAGAAGCGCTGAGAGACAGTCACAGTC 1206

39 SerCysLysAlaSerGlyTyrProPheInAsnIleCysIleMetAsnTrpValIleGlnAla 58
1207 TCTTCGAGGCGCTCTGGGTATCTCTTCAACAAGTATGAAATGAACTGGGGTGAAGCGGCT 1266

59 ProGlyGingIvLeuIysTrpMetGlyTrpIleAsnThrSerThrglyGluSerThrPhe 78

Db 1267 CCAGGACGCGGTTAAAGTGATGGCGTGGATTAAACACCTCCACTGGAGAGTCAACATTT 1322

QY 79 AlaAspSpHeLySGIYArgPheAspPheSerLeuGLuThrSerAlaAsnThrAlaTy 98

Db 1327 GGTGATGACTTCAAGGACGGTTGACTTCTTTGGAAACCTTGCCACACATGCCAT 1386

99 LeuGlnIleAsnMetLysSerGluAspMetAlaIleTyrPheCysAlaArgTyrGlu 118

Db 1387 TTGCAGATCAACAACCTCAAAAGTGAGACATGGCTACATATTTCTGTGCAGATGGGAG 144

119 **Val**ThrHisGlyTyrValProlytyrPolylgInglYmThrValThrsValSerSerGly 138

139 [G]W[Q]VSeGd[Q]W[Q]VSeGd[Q]W[Q]VSeAent[eq]n[e]uThrG[n] 158
144 / gH1AACACG3C1ACG11C11AC19G3CCAMGGACACGG1CACCG1T1C1C1GGC 150

1507 GGTGGCGATTCGGTGGCGGTGGCTCCGACCGGTGGCGGTTCTGACATCCAGCTGACCCCG 1566

159 SerHisIysPheLeuSerThrIseValGlyAspArgValSerIleThrCysLysAlaSer 178

Db 1567 TCTCAAAATTCCTGTCCACTTCAGTAGGAGACAGGGTCAGCATCACCCTGCAGGCCACT 1626

QY 179 GlnAspValTyrAsnAlaValAlaIatTrpTyrGlnGlnIysProGlnGlnSerProIysIeu 198

Db 1627 CAGAGTGTATATGCTGTTCCTGGTTCACAGAAACGAGCAATCTCTAACTT 1686

199 LeuileYrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 218

Db 1687 CTGATTACTGGCAATCCTCCCGGTACACTGGAGTCCCTTCTGCTTCACTGGCAGTGGC 1746

219 **ergGlyProaspPheThrPheThrIleSerSerValGlnAlaGluaspPheValAlaValTyr** 238

DB 1 / 1C1GGGCGCGAATTCACCTTACCCAGCAGTGICAGGCGTGCAGACTTAT 1806

239 pbaCyrcG]nG]nH; sDbaaywqthrbDbaaywqthrbGcl]usGacrl]ymhcl'ust'awcl']t.t.]c 258

b7
255 mecygcnagmnmrmnmgmgnmfnefnemefysecvylmlysdscuvdlle 258

b8
1807 ttctgtcagcacacattttccactcatcttccgctcgsgctcgsgggacaaataggagaatc 1866

Ox 259 LysAlaLeuGluIleSerAsnSerValMetTyrPheSerSer 272

Db 1867 AAAGCTAGAGGATCTCCAGTGACAGACATTTTCAGCC 1908

RESULT 12

AK268366					
LOCUS	AR268366	1919 bp	DNA	linear	PAT 10-APR-2003
DEFINITION	Genomic DNA from rat				
ACCESSION	GenBank	U040033			

DEFINITION	sequence	34	LOC101928616	US	04506253
ACCESSION	AR268366				
VERSION	AR268366.1	GI:29698716			

KEYWORDS	1
SOURCE	Unknown.

ORGANISM Unknown.
Unclassified.

REFERENCE
1 (bases 1 to 1919)
Wels, W. and Fominaya, J.
AUTHORS

JOURNAL	Patent: US 6498233-A 34 24-DEC-2002;
TITLE	Nucleic acid transfer system
ABSTRACT	
LOCATION	Location/Availability

Source	Location/Qualifiers
1. 1919	/organism="unknown"

```

/mol_type="genomic DNA"

```

Alignment Scores:

Pred. No.:	2.82e-92	Length:	1919
Score:	1308.00	Matches:	245

Percent Similarity: 97.64% Conservative: 3


```
QY 159 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 218
Db 655 CTGATTTCATCGGCATCTCCCGTACCTGAGTCCTTCTCGCTTCACTGCGAGTGGC 714
QY 219 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 238
Db 715 TCTGGGCGCGATTTCATCTTCACTGAGTCAGTGTGAGAGCTGAGAGCTGCGAGTTTAT 774
QY 239 PheGlyGlnGlnIlePheArgThrProPheThrPheGlySerGlyThrIleLeuGlnIle 258
Db 775 TCTGTGACACAACTTTGTGACTCCATTCACGTTCTGCGTGGGACAAATTTGAGATC 834
QY 259 LysAlaLeuGlnIleSerAsnSerValMetTyrPheSerSerValValProValLeuGln 278
Db 835 AAAGCTCTAGAGCA----- 848
QY 279 LysValAsnSerThrThrThrLysProValLeuArgThrProSerProValHisProThr 298
Db 849 -----CCATCATCAACATCACTAG 869
QY 299 GlyThrSerGlnProGlnArgProGlnAspCysArgProArgGlySerValIleGlyThr 318
Db 870 GGG---CGGACCTGCGCGCGCTGAC-----CGCGACACAGGCTGCCACTGCGCGCT 920
QY 319 GlyLeuAspPheLeuGluAspProLysLeuCysTyrIleLeuAspGlyIleLeuPheIle 338
Db 921 GGAGACTTTCACCCGTCATCGCCAGCGCGCGCTGGGACAACTGAGAGCACTGCGGCTA 980
QY 339 TyrGlyValIle-IleThrAlaLeuTyrLeuArgAlaLysPheSerArgSerAlaGluThr 358
Db 981 TCCGGTGAAGCGCGTGTGCTGCTTACCTGCGCGCGCGACTGTCA----- 1026
QY 358 AlaAlaAsnLeuGlnAspProAsnGlnLeuTyrAsnGlnLeuAsnLeuGlnIleArgArgG1 378
Db 1027 -----TGGACCAAGGTCGACCA-GG1----- 1046
QY 378 uGluTyrAspValLeuGlnIleLysArgAlaArgAspProGluMetGlyGlyIleGlnG1 398
Db 1047 -----GATCGGACAGCCCTGGCCAGCCCGC 1072
QY 398 nArgArgArgAsnPro-GlnGluGlyValTyrAsnAlaLeuGlnIleLysAspIleMetAlaG 418
Db 1073 GCAGCGGCGGCGACTGGGCGACAGATCCCGAGCACCGGAGAGGCGCGTGTGCGCC 1132
QY 418 IuaIaTyrSerGluIleGlyThrIleLysGlyIleArgArgArgGlyLysGlyHisAsp 436
Db 1133 TGACCCCTGGCGCGCGGACGAGCCGCTTCCTCGGACAGGCAAGCGCAACG 1188

RESULT 14
LOCUS A57343 2793 bp DNA linear PAT 03-MAR-1998
DEFINITION Sequence 13 from Patent EP0739984.
ACCESSION A57343
VERSION A57343.1 GI:3713222
KEYWORDS
SOURCE
ORGANISM
1
other sequences; artificial sequences.
REFERENCE
1
Wels,W.D., Schmidt,M. and Groner,B.D.
AUTHORS Bivalent polypeptides containing at least two domains
JOURNAL Patent: EP 0739984-A 13 30-OCT-1996;
SAN TUMORFORSCHUNGS GMBH (DE)
FEATURES
location/Qualifiers
source
1..2793
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/clone="PMS238-5-225"
1..2679
/note="unnamed protein product"
/codon_start=1
/transl_table=1
```

```
/protein_id="CAA03444.1"
/db_xref="GI:3713223"
/translation="MKRTALALAVLAGPVAQADYKDDDKLHHHHHKKLOVOLOQ
SGPELKRGEITVKSCKASGYPTNYGMNVKQAPGGGLKMGMINSTSESTPADP
KGRDPSLETSSANTAYLOINNEKSEDAATFCARMEYHYGVYWGQGVTVSSGGG
GGGGGGGGGSDILOTHSKFLSTSGVDSVSTCKSKASODVNAVAVYOOKPOSPPL
LIYASASRYNGVPRFTSGSGSDPFTTISVQABDLAVYFCOHPTPTPGSGTLE
EIKALEHHHHHHLEGGSSALATVHQAHLLEFTRRRRORRCMEQECGYPORLYA
LYLARLSKMQVDVIRNALASFGSGDLSEALREDEQARLALTALAAESKPVKQ
TGNDENAGANADEKLOVQLGSGPGLVQPSQSLITCTGFSGLTNYGVHVRQSPK
GLEMLGVWGGNTDYNTPPTSRLSINKDKSKQVFKMNSLONDTALTYCARALTY
YDVEFAWGGQVTVDVSSGGGSGGSGSDILOTHSPVILSVSGRVSFSCRA
SOSIGTNIHWYQORTNGSPRLIKVASESISGIPSRSGSGSDPFTTISNVSSEDI
ADYVCOONMNPPTTFGACTLAKIKALEDPVPSGAPSDGALLRMYTGAPEPCD
GGVRSFSTRQNTWTVRRLIQRQLREERYVYVGHGTRLEAOSIVREGVARSD
LDATWRGFYTAGDPALAYGADQDEPARIRINGALRVYVPRSSLPGRVRSLTLA
APEAAGVERLIGHPLRLDAITGPEEBEGRLETLIGMPLARVTVIPSALPTDPBN
VGQDLDPSSIPDKEQALISALPDYASQPGKPRBDLK"

ORIGIN
Alignment Scores:
Pred. No.: 4,51e-92 Length: 2793
Score: 1308.00 Matches: 275
Percent Similarity: 71.19% Conservative: 24
Best Local Similarity: 65.48% Mismatches: 59
Query Match: 53.58% Indels: 64
DB: 6 Gaps: 5

US-09-596-774-6 (1-461) x A57343 (1-2793)
QY 19 GlnIleGlnLeuValGlnSerGlyProGluLeuLysLeuProGlyGluThrValIle 38
Db 115 CAGGTAAACCTGACGACGACTGCACTGCACTGAAGAAGCTGAGAGACGCAAGATC 174
QY 39 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValIleGlnAla 58
Db 175 TCTGCAAGAGCCTCTGGGATCTCTTCACAAATGATGATGAATGAGTGGGAGAGGCT 224
QY 59 ProGlyGlnIleLysIleTyrPheMetGlyTyrPheAsnThrSerThrGlyLysThrPhe 78
Db 235 CAGGAGAGGGGTTTAAAGTGATGGCTGATTAACCTCCACCTGAGAGTCAACATTT 294
QY 79 AlaAspAspPheLysGlyArgPheAspPheSerLeuGlnIleThrSerAlaAsnThrAlaTyr 98
Db 295 GCGATGACTTCAAGGAGCGTGTGACTTCTCTTGGAAACCTCTGCCAACAATGCTGCT 354
QY 99 LeuGlnIleAsnAsnLeuLysSerGluAspMetAlaThrTyrPheCysAlaArgTyrGlu 118
Db 355 TTGCAGATCAACAACCTCAAAAGTGAAGCATGGCTACATATTTCTGTCAAGATGGAG 414
QY 119 ValTyrHisGlyTyrValProTyrTyrGlnGlyIleThrThrValIleSerSerGly 138
Db 415 GTTTACCAAGGCTACGTTCTTCTGAGGACCAAGGACCAAGGTCACCTTCTCTGCG 474
QY 139 GlyIleGlySerGlyIleGlyIleGlySerGlyIleGlyIleSerAspIleGlnLeuThrGln 158
Db 475 GGTGGCGTTCTGTGGCGGTGGCTCCGCGGGTGGCGGTTCGACATTCAGCTGACCCAG 534
QY 159 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSer 178
Db 535 TCTCACAAAATTCCTGTCCACTTCAGTAGGACAGAGGTAGGATCACTCGCAAGGCACT 594
QY 179 GlnAspValTyrAsnAlaValAlaTyrPyrGlnGlnLysProGlyGlnIleSerProLysLeu 198
Db 595 CAGGATGTATATATCTGTTGCTGTGATCAACAGAAACAGAGCAATCTCTAATCTT 654
QY 199 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 218
Db 655 CTGATTTCATCGGCATCTCCCGTACCTGAGTCCTTCTCGCTTCACTGCGAGTGGC 714
QY 219 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 238
Db 715 TCTGGGCGCGATTTCATCTTCACTGAGTCAGTGTGAGAGCTGAGAGCTGCGAGTTTAT 774
```

QY 239 PheCysGlnGlnIlePheArgThrProPheThrPheGlySerGlyThrLysLeuGlnIle 258
|||
Db 775 TTCTGTACAGCAACATTTTCTGACTCATTCACGTTCCGTCGGGACAAATTGGAGATC 834
QY 239 LysAlaLeuGlnIleSerAsnSerValMetCysPheSerSerValProValLeuGln 278
835 AAAGCTTAGAGCA----- 848
QY 279 LysValAsnSerThrThrThrLysProValLeuArgThrProSerProValHisProThr 298
849 -----CCATCATTCACCATCCTGAGA 869
QY 299 GlyThrSerGlnProGlnArgProGlnArgCysArgProArgGlySerValLysGlyThr 318
870 GGG---CGGACGCTCGCGCCGCTGAC-----CGCGACCAAGGCTCGCACTCCGCT 920
QY 319 GlyLeuAspPheLeuGlnArgProLysLeuCysTyrLeuLeuAspGlyTyrLeuPheIle 338
921 GGAGACTTTCACCCGTCATCGCAGCGCGGCTGGGAAACAAGTGGAGCAGTGGGCTA 980
QY 339 TyrGlyValIle-IleThrAlaLeuTyrLeuArgAlaLysPheSerSerValGluThr 358
981 TCCGGTGACGCGCTGCTGCTCTTACCTGCGCGCGGCACTGCA----- 1026
Db 358 AlaAlaAsnLeuGlnArgProAsnGlnLeuTyrAsnGlnLeuAsnLeuGlyArgArgGly 378
1027 -----TGGAACCAAGTCCGCA-GGT----- 1046
QY 378 uGluTyrAspValLeuGlnLysLysArgAlaArgAspProGlnMetGlyGlyLysGlnI 398
1047 -----GATCCGCAAGCCCGGCCCCCG 1072
QY 398 nArgArgArgAsnPro-GlnGlnGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaG 418
Db 1073 GCAGCGCGGCGACCTGGCGGAGGATCCGCGACAGCGCGGAGCGCGCTCTGGGCC 1132
QY 418 LysAlaTyrSerGlnIleGlyThrLysGlyGlyLysArgArgGlyLysGlyHisAsp 436
Db 1133 TGACCCCTGCGCGCGCGGACGAGCGGCTTCTCGCGGAGGCGACCGGCAAGCAGC 1188

RESULT 15
LOCUS A57351 2793 bp DNA linear PAT 03-MAR-1998
DEFINITION Sequence 21 from Patent EP0739984.
ACCESSION A57351
VERSION A57351.1 GI:3713230
KEYWORDS
SOURCE
ORGANISM
other sequences; artificial sequences.
REFERENCE
1 Wels,W.D., Schmidt,M. and Groner,B.D.
AUTHORS Bivalent polypeptides containing at least two domains
TITLE Patent: EP 0739984-A 21 30-OCT-1996;
JOURNAL SAN TIMORORSCHINGS GMBH (DE)
FEATURES
SOURCE
1. 2793 Location/Qualifiers
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:3630"
/clone="PMS238-5-5"
/note="MUSE; P.ABRUGINOSA; SYNTHETIC"
1. 2679
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAA03448.1"
/db_xref="GI:3713231"
/translation="MKTAIAIAVALAFATVAADYKDDDKLHHHHHKLQVLOO
SGPELKKPGETVKISCKASGYPTNYGNWKKQAGGLKMMGMINTSGSTPADR
KGRPFSLSTANTAYLQIINLKSEDMATYFCARWEYVHGVPVWGQCTVTVSSGG
GSGGGSGGSGDIDLTQSHKFLSTSVDRVSICTKASGDVYNAVAYQKPGSPKL

ORIGIN
Alignment Scores:
Pred. No.: 4,51e-92 Length: 2793
Score: 1308.00 Matches: 275
Percent Similarity: 71.19% Conservative: 24
Best Local Similarity: 65.48% Mismatches: 59
Query Match: 53.58% Indels: 64
DB: 6 Gaps: 5
US-09-596-774-6 (1-461) x A57351 (1-2793)
QY 19 GlnIleGlnLeuValGlnSerGlyProGlnLeuLysLysProGlyGluThrValLysIle 38
Db 115 CAGGTACACCTCGACGACACTCGACCTGAACTGAAGAAGCCTGGAGAGACAGTCAAGTTC 174
QY 39 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValLysGlnAla 58
Db 175 TCTGTCAAGGCGCTCGGGTATCTTTTCAAACTATGGAAGAACTGGGTGAAGCGGCT 234
QY 59 ProGlyGlnGlyLeuLysTyrMetGlyTyrIleAsnThrSerThrGlyLysSerThrPhe 78
Db 235 CCAGGACAGGGGTTTAAAGTGAATGGATGGCTGATTAACCTCAGTGGAGAGTCAACATT 294
QY 79 AlaAspAspPheLysGlyArgPheAspPheSerLeuGlnThrSerAlaAsnThrAlaTyr 98
Db 295 GGTATGACTTCAAGGAGCGGTTTGACTCTCTTTGGAACCTCGCCAACTGCTCCAT 354
QY 99 LeuGlnIleAsnAsnLeuLysSerGlnAspMetAlaThrTyrPheCysAlaArgTyrGlu 118
Db 355 TTGCAGATCAACAACCTCAAAAGTGAAGACATGCTACATTTTCTGTGCAAGATGGAG 414
QY 119 ValTyrHisGlyTyrValProTyrTyrGlnGlnGlyThrThrValThrValSerSerGly 138
Db 415 GTTACCAAGGCTACGTTCTTACTGGGGCCAAAGGAGCACAGGTACCGTTCTCTGGC 474
QY 139 GlyGlyGlySerGlyGlyGlySerGlyGlyGlyGlySerAspIleGlnLeuThrGln 158
Db 475 GGTGGCGGTTTGTTGGCGGTGGCTCCGGCGGTGGCTTGTGATCCACTGACCTGAC 534
QY 159 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysAlaLysSer 178
Db 535 TCTCACAAAATTCCTGTCACCTCAGTAGAGACAGGCTCACATCAGTCAAGGCCAGT 594
QY 179 GlnAspValTyrAsnAlaValAlaTyrTyrGlnGlnLysProGlyGlnSerProLysLeu 198
Db 595 CAGATGTGTAATATGCTGTGCTGTGATCAACAGAAACAGAGCAATTCCTTAACCT 654
QY 199 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 218
Db 655 CTGATTACTGGGATCTCCCGGTACACTGGAATCTCTTCCGTTCACTGCGCAGTGGC 714
QY 219 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGlnAspLeuAlaValTyr 238
Db 715 TCTGGGCGGATTTCACTTCAACATCAGCAGTGTGAGCGCTGAAGACTGGCAGTTAT 774
QY 239 PheCysGlnGlnIlePheArgThrProPheThrPheGlySerGlyThrLysLeuGlnIle 258
Db 775 TTCTGTACAGCAACATTTTCTGACTCATTCACGTTCCGTCGGGACAAATTGGAGATC 834
QY 259 LysAlaLeuGlnIleSerAsnSerValMetCysPheSerSerValProValLeuGln 278

```
Db      835 |||||
      835 AAAGCTTAGAGCA----- 848
QY      279 LysValAsnSerThrThrLysProValLeuArgThrProSerProValHisProThr 298
      849 -----CCATCATCACCATCAGCTAG 869
QY      299 GlyThrSerGlnProGlnArgProGlnAspCysArgProArgGlySerValLysGlyThr 318
      870 GGG---CGGCAAGCTGGCGCGCTGAC-----CGGCAAGCTGGCGCGCTGACCTGCCCT 920
QY      319 GlyLeuAspPheLeuGlnAspProLysLeuCysTyrLeuLeuAspGlyTleuPheIle 338
      921 GGAGACTTTCACCCCTCATCGCCAGCGCGCTGGGAACTGAGCAAGTGCAGCTA 980
QY      339 TyrGlyValIle-IleThrAlaLeuTyrLeuArgAlaLysPheSerArgSerAlaGluTh 358
      981 TCCGGTGCAGCGGCTGCTGCTCTTACTGCGCGCGCTGCA----- 1026
QY      358 rAlaAlaAsnLeuGlnAspProAsnGlnLeuTyrAsnGlnLeuAsnLeuGlyArgArgG 378
      1027 -----TGGAACCAAGTCAGCA-GGT----- 1046
QY      378 uGluTyrAspValLeuGlnLysLysArgAlaArgAspProGluMetGlyGlyLysGlnG 398
      1047 -----GATCCGCAAGCCCTGCGCAGCCCG 1072
QY      398 nArgArgArgAsnPro-GlnGlnGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaG 418
      1073 GCAGCGGCGCGAGCTGGCGGAGATCCGCGAGCAGCGGAGCAGGCCGCTGCGCTC 1132
QY      418 IuaIaTyrSerGlnIleGlyThrLysGlyGluArgArgGlyLysGlyHisAsp 436
      1133 TGACCCCTGGCGCGCGAGAGCGAGCGCTTCTGCGGAGGAGCGGCAACGAC 1188
Db
```

Search completed: April 20, 2005, 14:06:06
Job time : 6200.38 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Comphen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 20, 2005, 07:10:49 ; Search time 735.356 Seconds
(without alignments)
3711.128 Million cell updates/sec

Title: US-09-596-774-6

Perfect score: 2441
Sequence: 1 MAMVTLFLMLAAKVPKQI.....ISTATKDTYDALHMQLAPR 461

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

MODEL=frame+g2n.model -DEV-xlh
-O=/cgn2.1/USPTO_epool_h/US09596774/runat_20042005_075441_12501/app_query.fasta.1.1230
-DB=N_Geneseq -QFM=fastap -SUFFIX=ap20.ing -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=biosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTPM=PCO -NOR=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09596774_QCGN_1.1.653@runat_20042005_075441_12501 -NCPU=6 -ICPU=3
-NO_MMAP -LARGQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOC
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseq_16dec04:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2441	100.0	1479	2	AAT05783 Leader-sc
2	1675	68.6	1452	2	AAV73324 Chimeric
3	1506.5	61.7	3058	2	AAV77137 Single ch
4	1503.5	61.6	1284	2	AAV77139 Single ch
5	1314.5	53.9	2011	2	AAQ28262 Fv(FRPs) -

6	1312.5	53.8	2070	2	AAT42036	Aat42036 Plasmid p
7	1312.5	53.8	2730	2	AAT42044	Aat42044 Plasmid p
8	1308	53.6	1862	2	AAT29413	Aat29413 DGA14-DT-
9	1308	53.6	1919	2	AAT29412	Aat29412 DGA14-DT-
10	1308	53.6	2214	2	AAT42038	Aat42038 Plasmid p
11	1308	53.6	2793	2	AAT42039	Aat42039 Plasmid p
12	1308	53.6	2793	2	AAT42043	Aat42043 Plasmid p
13	1308	53.6	3177	2	AAT42041	Aat42041 Plasmid p
14	1305	53.5	1692	2	AAT29409	Aat29409 scFv(FRPs)
15	1303	53.4	2233	2	AAQ28257	AaQ28257 Fv(FRPs) -
16	1302	53.3	794	2	AAT42034	Aat42034 Plasmid p
17	1302	53.3	2718	2	AAT42042	Aat42042 Plasmid p
18	1302	53.3	2793	2	AAT42040	Aat42040 Plasmid p
19	1285.5	52.7	748	2	AAT05781	Aat05781 ErB-2-EP
20	1276.5	52.3	748	2	AAQ28256	AaQ28256 Fv heavy
21	1249	51.2	1920	10	ADH34695	Adh34695 CE7-speci
22	1249	51.2	6834	10	ADH34699	Adh34699 CE7-speci
23	1015.5	41.6	7654	8	ABX16565	Abx16565 Retrovira
24	1015.5	41.6	7654	10	ABX13168	Abx13168 Retrovira
25	1013	41.5	726	4	AAK30723	Aak30723 Anti-FIX/
26	1011.5	41.4	1422	2	AAT90509	Aat90509 DNA encod
27	1011.5	41.4	1545	2	AAT90510	Aat90510 DNA encod
28	994	40.7	771	4	AAC86584	Aac86584 DNA encod
29	994	40.7	771	9	ABT43523	Abt43523 scFvSA fu
30	981.5	40.2	909	2	AAT36878	Aat36878 741F8 ant
31	981.5	40.2	909	2	AAT91837	Aat91837 Anti-c-er
32	981.5	40.2	909	2	AAT78879	Aat78879 Single ch
33	981.5	40.2	909	2	AAV22598	Aav22598 DNA encod
34	981.5	40.2	909	2	AAV21796	Aav21796 741F8 ant
35	981.5	40.2	909	2	AAV63397	Aav63397 Anti-c-er
36	981.5	40.2	909	10	AAD61483	Aad61483 741F8 gfv
37	978.5	40.1	987	10	ADG32348	Adg32348 DNA encod
38	972.5	39.8	777	10	ADG32297	Adg32297 Mouse scF
39	972.5	39.8	2001	2	AAQ46088	AaQ46088 Sequence
40	963.5	39.5	987	10	ADG32350	Adg32350 DNA encod
41	962.5	39.4	1956	2	AAT90512	Aat90512 DNA encod
42	962.5	39.4	2079	2	AAT90513	Aat90513 DNA encod
43	960.5	39.3	777	10	ADG32299	Adg32299 Mouse scF
44	930	38.1	1350	3	AAT97844	Aat97844 DNA of a
45	928.5	38.0	800	3	AAA08034	Aaa08034 Antibody

ALIGNMENTS

RESULT 1	
ID AAT05783	standard; cDNA, 1479 BP.
XX	
AC AAT05783;	
XX	
DT 16-MAR-1996	(first entry)
XX	
DE Leader-scFv (FRPs): IYc-2 hinge:zeta cDNA.	
XX	
KW Single chain antibody; scFv; antibody engineering; antitumour;	
KW tumour antigen binding; T-cell receptor; cytotoxic T-lymphocyte;	
KW monoclonal antibody; erbB-2; cancer; cell targeting; gene therapy;	
KW adoptive immunotherapy; ss.	
XX	
OS Synthetic.	
XX	
FT Key	Location/Qualifiers
FT CDS	40..1426
FT	/*tag= a
FT	/product= "Leader-scFv(FRPs): IYc-2 hinge:zeta"
XX	
PN W09530014-A1.	
XX	
PD 09-NOV-1995.	
XX	
PF 20-APR-1995;	95WO-EP001494.
XX	
PR 02-MAY-1994;	94EP-00810244.

XX. (CIBA) CIBA GEIGY AG.
 PA Groner B, Moritz D;
 PI WPI; 1995-393085/50.
 DR P-PSDB; AAR85508.
 XX
 PT New bifunctional proteins for use in killing tumour cells - contg. a
 PT tumour antigen binding domain, a hinge region and a zeta chain derived
 PT from a T-cell antigen receptor.
 PS Example 1; Page 29-31; 46pp; English.
 XX cDNA (AAT05783) codes for a bifunctional protein (AAR85505) consisting of
 CC single chain antibody scFv(FRP)5 directed against the tumour erbB-2
 CC antigen, a hinge region, and a functional zeta chain obd. from a T-cell
 CC receptor. The cDNA is expressed in host cells, esp. cytotoxic T-
 CC lymphocytes, providing them with a defined tumour cell specificity
 CC enabling targeting to defined tumour cells and MHC-unrestricted and MHC-
 CC independent tumour destruction in vitro or in vivo
 XX
 SQ Sequence 1479 BP; 369 A; 395 C; 394 G; 321 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 3,02e-183 Length: 1479
 Score: 2441.00 Matches: 461
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: Gaps: 0
 US-09-596-774-6 (1-461) x AAT05783 (1-1479)
 QY 1 MetAlaTTPvAlTTPThLeuLeuPheLeuMeAlaAlaAlaAlaValProLySGlnIle 20
 DB 40 ATGGCTTGGGTGGATCCCTGCTATCTCGATGGAGCTGCCAAGTCCCAAGAGATC 99
 QY 21 GlnLeuValGlnSerGlyProGlnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 40
 DB 100 CAGTTGGTGGAGTCTGGAGCTGGAGTGAAGAGCTGGAGAGACAGTCAATCTCTGCTGC 159
 QY 41 LysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTTPvAlLysGlnAlaProGly 60
 DB 160 AAGGCTCTGGGTATCTCTTTCACAAATATGAAATGAACTGGGTGAAGCAGCTCCAGA 219
 QY 61 GlnGlyLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 80
 DB 220 CAGGCTTTAAAGTGGATGGCTGGATTAACACCTTCACCTGGAGAGTCAACATTTGCTGAT 279
 QY 81 AspPheLeuGlyLysArgPheAspPheSerLeuGlnThrSerAlaAsnThrAlaTyrLeuGln 100
 DB 280 GACTTCAAGGAGCGGTTGACTTCTCTTGGAAACCTCTGCCAACATCGCTTATTTGACG 339
 QY 101 IleAsnLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 120
 DB 340 ATCAACAACCTCAAAAGAGACATGCTCATATTTCTGCAAGATGGAGAGTTTAC 399
 QY 121 HisGlyTyrValProTyrTTPGlyGlnGlyThrThrValThrValSerSerGlyGlyGly 140
 DB 400 CACGGCTTACGTTCTTACTTGGGCGCAAGAGACACCGCTCACCGTTCTCTGGCGGTGGC 459
 QY 141 GlySerGlyGlyGlyGlySerGlyGlyGlyGlyGlySerAspIleGlnLeuThrGlnSerHis 160
 DB 460 GGTTCTGTGGGTGGGTCCGGCGGTGGCGGTCTGACATCCAGTCAACCGAGTCTCAC 519
 QY 161 LysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSerGlnAsp 180
 DB 520 AAATTCCTGTCCACTTCAAGAGAGACAGGCTCAGCTACCTCCCAAGGCGCAGCAGAT 579
 QY 181 ValTyrAsnAlaValAlaTTPYrGlnGlnLysProGlyGlnSerProLysLeuLeuIle 200
 DB 580 GTGATATATGCTGTGCTGGTATCAACAGAAACAGAGCAATCTCTTAACATTCTGATT 639

QY 201 TyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 220
 DB 640 TACTGGCATCTCCCGGTACATGAGTCCCTTCTGCTTCACTGGCAGTGGCTCTGGG 699
 QY 221 ProAspPheThrPheThrIleSerSerValGlnAlaGlnAspLeuAlaValTyrPheCys 240
 DB 700 CCGGATTCACTTTCACCATCAGATGAGTGGAGCTGAGAGCTGGAGAGTTATTTCTGT 759
 QY 241 GlnGlnHisPheArgThrProPheThrPheGlySerGlyThrLysLeuGlnLysAla 260
 DB 760 CAGCAACATTTCTGACTCCATTCACCTGGCTGGGCGCAAAATTTGGAGTCAAAAGCT 819
 QY 261 LeuGlnIleSerAsnSerValMetTyrPheSerSerValValProValLeuGlnLysVal 280
 DB 820 CTAAAGATCAACAACCTGGATGATGACTTCAATGTTCTGTCTGCGCAAGTCTTCAGAAAGTG 879
 QY 281 AsnSerThrThrThrLysProValLeuArgThrProSerProValHisProThrGlyTyr 300
 DB 880 AACTTCTACTACTACCAAGCCAGTCTGCAATCTCTCACTGTGCACTTACCGGAGCA 939
 QY 301 SerGlnProGlnArgProGlnAspCysArgProArgGlySerValLysGlyThrGlyLeu 320
 DB 940 TCTCAGCCCGCAGAGACAGAAAGATTGTGCGCCCGTGTCTCAGTGAAGGGGACCGGATTG 999
 QY 321 AspPheLeuGlnLysArgProLysLeuCysTyrLeuLeuAspGlyLysLeuPheIleTyrGly 340
 DB 1000 GACTTCTTAAAGATGCCCAAACTCTGCTACTCTGATGAGAAATCTCTTCACTTACCGGA 1059
 QY 341 ValIleIleThrAlaLeuTyrTyrLeuArgAlaLysPheSerArgSerAlaGlnThrAlaAla 360
 DB 1060 GTCATATCAACAGCCCTGTACTCTGAGAGCAAAATTCAGCAGAGTGCAGAGACTCTGCC 1119
 QY 361 AsnLeuGlnAspProAsnGlnLeuTyrAsnGlnLeuAsnLeuGlnLysArgGlnGlyTyr 380
 DB 1120 AACCTGAGAGACCCCAACAGCTCTCAATGAGCTCAATCTAAGGAGAGAGAGAAATAT 1179
 QY 381 AspValLeuGlnLysLysAlaArgAlaArgAspProGlnMetGlyGlyLysGlnArgArg 400
 DB 1180 GAGCTCTTGAAGAAAGAGCGGGCTCGGATTCAGAGATGGAGGCAAAACAGCAGAGAGG 1239
 QY 401 ArgAsnProGlnGlnGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaGlnAlaTyr 420
 DB 1240 AGGAACCCCGAGAGAGCGGTATCAATGCACTGCAGAAAGACAAGATGGAGAAAGCTTAC 1299
 QY 421 SerGlnIleGlyThrLysGlyLysArgArgArgGlyLysGlyHisAspGlyLeuTyrGln 440
 DB 1300 AGTGAATTCGGCAAAAGGCGAGGCGAGGAGGCAAGGGGCAAGATGGCTTTTACAG 1359
 QY 441 GlyLeuSerThrAlaThrLysAspThrTyrAspAlaLeuHisMetGlnThrLeuAlaPro 460
 DB 1360 GGTCTCAGACATCCCAACCAAGACACCTATGATGCCCTCATATGCAAGACCTCGCCCT 1419
 QY 461 Arg 461
 DB 1420 CGC 1422
 RESULT 2
 ID AAV73324 standard; DNA; 1452 BP.
 XX
 AC AAV73324;
 XX
 DT 26-FEB-1999 (first entry)
 PT
 DE Chimeric CD19/CTCR DNA.
 XX
 XX Single chain; antibody; B-cell marker; CD19; scFvCD19; gene therapy;
 KM retroviral vector; tumour-associated antigen; cancer; immunoglobulin;
 KM CD8 alpha chain; hinge region; theta chain; transmembrane domain;
 KM intracellular domain; T-cell receptor-CD3 complex; ss.
 XX
 OS Homo sapiens.

OY 459 ALAProArg 461
 DB 1441 GCCCCTCGC 1449

RESULT 3
 AAT77137
 ID AAT77137 standard; DNA; 3058 BP.
 AC AAT77137;
 XX
 XX 25-MAR-2003 (revised)
 DT 04-MAR-1998 (first entry)
 DE Single chain antigen hybrid receptor DNA.
 XX
 XX Hybrid receptor; single chain antigen; gene therapy; diagnosis;
 KW signal conduction; receptor; control region; ss.
 OS Synthetic.
 OS Homo sapiens.

Key Location/Qualifiers
 FH misc_signal 562..570
 FT /tag= a
 FT /note= "Noted in specification"
 FT CDS 723..2136
 FT /tag= b
 FT /note= "This sequence contains an intron"
 FT sig_peptide 723..861
 FT /tag= d
 FT /label= leader sequence
 FT /note= "This sequence contains an intron"
 FT intron 768..849
 FT /tag= c
 FT mat_peptide 862..2133
 FT /tag= e
 FT /product= "single_chain_antigen_hybrid_receptor"
 FT polya_site 2514..2519
 FT /tag= f

MO9720938-A2.
 XX 12-JUN-1997.
 PD
 XX
 XX 03-DEC-1996; 96WO-DE002334.
 PE
 XX 05-DEC-1995; 95DE-01045351.
 PR
 XX (UYFR-) UNIV FREIBURG KLINIKUM ALBERT-LUDWIGS.
 PA
 XX Mertelsmann R, Kulmburg P, Rosenthal F;
 PI MPI; 1997-319784/29.
 DR P-PSDB; AAW24025.
 XX
 XX Cells with hybrid receptor having extracellular and intracellular regions
 PT of different origins - useful in gene therapy and diagnosis of tumours.
 FT
 XX
 XX Example 3; Fig 4; 46pp; German.

This DNA sequence encodes a novel single chain antigen hybrid receptor (HR) and contains an extracellular domain specific for the hapten 4-hydroxy-5-iodo-3-nitrophenyl acetate (NIP) coupled to the CD8-alpha region and the transmembrane and signal-conducting intracellular parts of the CD3-zeta molecule. Such hybrid receptors comprise a receptor part localised on the outside of the cell and specific to a particular signal molecule and a receptor part originating from another receptor, localised on the inside of the cell and capable of setting off a signal inside the cell. The cell should also contain at least one other gene construct with a control region which can interact with the signal sent out by the hybrid receptor and thereby control expression of a transgene bound to this control region. Such cells are useful in gene therapy or for

CC diagnostic purposes. (Updated on 25-MAR-2003 to correct PR field.)
 XX
 SQ Sequence 3058 BP; 689 A; 831 C; 853 G; 685 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 4,43e-109 Length: 3058
 Score: 1505.50 Matches: 299
 Percent Similarity: 75.27% Conservative: 45
 Best Local Similarity: 65.43% Mismatches: 86
 Query Match: 61.72% Indels: 27
 DB: 2 Gaps: 6

US-09-596-774-6 (1-461) x AAT77137 (1-3058)

OY 9 PheLeuMetAlaAlaAlaValProLysGlnIleGlnLeuValGlnSerGlyProGlu 28
 DB 832 TTTCCTCTTCTCTCCACAGGTGTCACCTCCAGGTCACCTGACGACGTCGTGGGCTGAG 891

OY 29 LeuLysLysProGlyGlnThrValLysIleSerCysLysAlaSerGlyTyrProPheThr 48
 DB 892 CTTGTGAAGCCTGGGGCTTCAGTGAACCTGCTCCGCAAGCCTTCTGCTACCTTACC 951

OY 49 AsnTyrGlyMetAsnTyrValLysGlnAlaProGlyGlnIleuLysTyrMetGlyTyr 68
 DB 952 AGCTACTGATGCACTGGGTGAAGCAAGGCTGSAAGGCTTGATGATTTGAAGG 1011

OY 69 IleAsnThrSerThrGlyGlnSerThrPheAlaAspPheLysGlyArgPheAspPhe 88
 DB 1012 ATTGATCTTAATGTTGGTGGTACTCAAGTACATGAGAGTTCAGAGCAAGCCACACTG 1071

OY 89 SerLeuGlnThrSerAlaAsnThrAlaTyrLeuGlnIleAsnAsnLeuLysSerGluAsp 108
 DB 1072 ACTGTAGCAAAACCTCCACACAGGCTTCATGACGTACAGCTGACGCTCATCTGAGAC 1131

OY 109 MetAlaThrTyrPheCysAlaArgTyrGluValTyr---HisGlyTyrValProTyrTyr 127
 DB 1132 TCTGGGCTCATATGTTATGTCMAAGATGATGATTACTAGGATGATCTTGTGACTCTGG 1191

OY 128 GlyGlnGlyThrThrValThrValSerSerGlyGlyGlyGlySerGlyGlyGlySer 147
 DB 1192 GGCACAGGACCCAGGTACCGCTCTCTCAGGGAGGGAGGTTCAGGAGGAGGTGCTCT 1251

OY 148 GlyGlyGlyGlySerAspIleGlnLeuThrGlnSerHisLysPheLeuSerThrSerVal 167
 DB 1252 GCGCGTGGCGGATCTCAGGCTGTGGACACAG---GAATCTGCATCTCACCATCACT 1308

OY 168 GlyAspArgValSerIleThrCysLysAlaSerGlnAspValTyr-----AsnAla 184
 DB 1309 GTGAAACAGTCACACTCACTGCTGCTCAAGTACGAGGCTGTTCACACTAGTAACTAT 1368

OY 185 ValAlaTyrTyrGlnGlnLysProGlyGlnSerProLysLeuLeuIleTyrSerAlaSer 204
 DB 1369 GCCAAGTGGGTCCAAAGAAAACCAATCATTTATTACAGTGTCTAATAGGTGTACCAAC 1428

OY 205 SerArgTyrThrGlyValProSerArgPheThrGlySerGlySerGlyProAspPheThr 224
 DB 1429 AACCGAGCTCCAGGTGCTCTGCTGCAAGATTCTCAGGCTCCCTGATGACACAAGGCTGC 1488

OY 225 PheThrIleSerSerValGlnAlaGlnAspLeuAlaValTyrPheCysGlnIleHisPhe 244
 DB 1489 CTCACATCAACAGGGCCACAGTGAAGTAGAGCAATATTTCTGCTCTATAGTAC 1548

OY 245 ArgThrProPheThrPheGlySerGlyThrLysLeuGlnIleLysAlaLeuGlnIleSer 264
 DB 1549 AGCAACCATTTGGTGTTCGTGGAGGAACAATGACATGACGTC-----CTCAG----- 1596

OY 265 AsnSerValMetCysThrPheSerSerValValProValLeuGlnLysValAsnSerThrThr 284
 DB 1597 -----GAATCTACTACT 1608

OY 285 ThrLysProValLeuArgThrProSerProValHisProThrGlyThrSerGlnProGln 304
 DB 1609 ACCAAGCCAGTCTGGAATCTCTCACTGTCACCTTACCGGAGACATCTCAGCCCCAG 1668

```
OY 305 ArgProGluAspGlySerProArgGlySerValIleGlyThrGlyLeuAspPheLeuGlu 324
DB 1669 AGACCAAGAGATTGCGCCCTGGCTCAGTAAAGGGACCGGATTGGACCTTC----- 1722
OY 335 AspProGlyLeuGlySerLeuAspGlyIleLeuPheIleTyrGlyValIleIleThr 344
DB 1723 GACCCCAAACTCTGCTACTGCTAGATGGAATCCTTTCATCTACGAGTCAATCATCA 1782
OY 345 AlaLeuTyrLeuArgAlaIlePheSerArgSerAlaGluThrAlaAlaAsnLeuGlnAsp 364
DB 1783 GCCCTGTACTCTGAGACGAATAATTCAAGCAAGTGCAGACATGCTGCCAACTCGACGAGAC 1842
OY 365 ProAsnGlnLeuTyrAsnGlnLeuAsnLeuGlnIleArgArgGlnGluIleTyrAspValLeuGlu 384
DB 1843 CCCAACCAAGCTCTACAAATGAGCTCAATCAAGGCGCAAGAGAAATATGACCTCTTGAG 1902
OY 385 LysIleAspAlaArgAspProGlnMetGlyIleGlnGlnArgArgAsnProGln 404
DB 1903 AAGAAAGCGGGCTCGGATCCAGATGGGAGCCAAACGACAGAGAGAGAGAACCCCGAC 1962
OY 405 GluGlyValIleTyrAsnAlaLeuGlnIleLysAspIleMetAlaGluAlaTyrSerGluIleGly 424
DB 1963 GAAGGCATATCAATGCACTGCGAAGAACCAACATGCGAGAGAGCTTACAGTGAATCGGC 2022
OY 425 ThrIleGlyIleArgArgGlyIleGlyIleAspGlyLeuTyrGlnGlyLeuSerThr 444
DB 2023 ACAAAGGCGAGAGGCGGAGAGGAGGAGGAGGAGTGGCTTTTACAGAGGCTTCAGCACT 2082
OY 445 AlaThrIleAspThrTyrAspAlaLeuHisMetGlnThrLeuAlaProArg 461
DB 2083 GCCACCAAGACACCTATGATGCTCCCTGCATATGACGACCCCTGCGC 2133

RESULT 4
AAT77139
ID AAT77139 standard; DNA; 1284 BP.
XX
AC AAT77139;
XX
DT 25-MAR-2003 (revised)
DT 04-MAR-1998 (first entry)
XX
DE Single chain antigen hybrid receptor partial DNA sequence.
XX
KM Hybrid receptor; single chain antigen; gene therapy; diagnosis;
KM signal conduction; receptor; control region; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..1276
FT FT /*tag= a
FT FT /note= "partial hybrid receptor"
FT FT sig_peptide 1..12
FT FT /*tag= b
FT FT /note= "partial leader sequence"
FT FT mat_peptide 13..1285
FT FT /*tag= c
FT FT /product= "hybrid receptor"
FT FT /note= "single chain antigen hybrid receptor"
XX
PN M09720938-A2.
XX
PD 12-JUN-1997.
XX
PF 03-DEC-1996; 96MO-DE002334.
XX
PR 05-DEC-1995; 95DE-01045351.
XX
PA (UYFR-) UNIV FREIBURG KLINIKUM ALBERT-LUDWIGS.
XX
PI Mettelmann R, Kulmburg P, Rosenthal F;
```

```
XX MPI, 1997-319784/29.
DR P-PSDB; AAM24027.
XX
PT Cells with hybrid receptor having extracellular and intracellular regions
PT of different origine - useful in gene therapy and diagnosis of tumours.
XX
PS Example 3; Fig 4; 46pp; German.
XX
CC This partial DNA sequence encodes a novel single chain antigen hybrid
CC receptor (HR). The full length receptor sequence can be found in AAT77137
CC and contains an extracellular domain specific for the hapten 4-hydroxy-5-
CC iodo-3-nitrophenyl acetate (NIP) coupled to the CD8-alpha region and the
CC transmembrane and signal-conducting intracellular parts of the CD3-zeta
CC molecule. Such hybrid receptors comprise a receptor part localised on the
CC outside of the cell and specific to a particular signal molecule and a
CC receptor part originating from another receptor, localised on the inside
CC of the cell and capable of setting off a signal inside the cell. The cell
CC should also contain at least one other gene construct with a control
CC region which can interact with the signal sent out by the hybrid receptor
CC and thereby control expression of a transgene bound to this control
CC region. Such cells are useful in gene therapy or for diagnostic purposes.
CC (Updated on 25-MAR-2003 to correct PR field.)
XX
SQ Sequence 1284 BP; 334 A; 342 C; 355 G; 253 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 2.68e-109 Length: 1284
Score: 1503.50 Matches: 298
Percent Similarity: 76.00% Conservative: 44
Best Local Similarity: 66.22% Mismatches: 81
Query Match: 61.59% Indels: 27
DB: Gaps: 6

US-09-596-774-6 (1-461) x AAT77139 (1-1284)
OY 16 ValProGlyGlnIleGlnLeuValGlnSerGlyProGlnLeuLysProGlyGlnThr 35
DB 4 GTCCACTCCCAAGTCCAACTGCAGACAGCTGGGGCTGAGCTTGGAAACCTGGGGCTTCA 63
OY 36 ValIleSerGlyValIleSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrVal 55
DB 64 GTGAAGCTGTCTCCAAAGGCTTGGCTGCTACCTTCACTCACTGATGATCAGCTGGCTG 123
OY 56 LysGlnAlaProGlyGlnGlyLeuLysTrpMetGlyTyrPheAsnThrSerThrGlyGlu 75
DB 124 AAGCAGAGGCTGAGCGAGGCTTGAAGATTGATCTCAATATGATGTGCT 183
OY 76 SerThrPheAlaAspAspPheLysGlyIleArgPheAspPheSerLeuGluThrSerAlaAsn 95
DB 184 ACTAAGTCAATGAGAACTTCAAGACAGGACACATGATGATGACAAACCTCCAC 243
OY 96 ThrAlaTyrLeuGlnIleAsnAlaLeuLysSerGlnAspMetAlaThrTyrPheCysAla 115
DB 244 ACAGCTTACATGACAGCTCAGCAGCTGATGAGTGTGAGCTGCGGTATTAATGTGCA 303
OY 116 ArgTrpGluValTyr---HisGlyTyrValProTyrTrpGlyGlnGlyThrValThr 134
DB 304 AGATACGATTACTACGAGTGAAGTACTTGAATCACTGGGGCCAAAGGACACAGCTCAC 363
OY 135 ValSerSerGlyGlyIleGlySerGlyIleGlySerGlyIleGlySerAspIle 154
DB 364 GTTCCTCAGGTGAGGCGGCTTCAAGCGGAGTGGCTTGGCGGTGCGGATCTCAGGCT 423
OY 155 GlnLeuThrGlnSerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThr 174
DB 424 GTTGAGACACAG---GAATGCACTCACACATCACTGCTGTAACAGTCACTCACT 480
OY 175 CysIleValIleSerGlnAspValTyr-----AspAlaValAlaTyrPyrGlnGlnLys 191
DB 481 TGTGGCTCAAGTACTGGGGCTGTAAACATGATTAATGCAACATGGGCTCAAGAAATA 540
OY 192 ProGlyGlnSerProLysLeuLeuIleTyrSerAlaSerSerArgTyrThrGlyValPro 211
```

```

Db      541 CCAGATCAATTATTTACTGCTGCTAATAGGNGTACACACACGAGCTCCAGTGTCTCT 600
Qy      212 SerArgPheThrGlySerGlySerGlyProAspPheThrPheThrLysSerValGln 231
Db      601 GCGAATTTCTCGGCTCCCTGATTGGAGACAAAGCTCCCTCCCTACCATCAAGGCGACAG 660
Qy      232 AlaGluAspLeuAlaValTyrPheCysGlnGlnHisPheArgThrProPheThrPheGly 251
Db      661 ACTGAGAGAGAGGACATATATTTCTGTCTGTATGTTACAGCAACCATTTGGGTGTTGGT 720
Qy      252 SerGlyThrLysLeuGlnLysAlaLeuGlnLysSerAsnSerValMetTyrPheSer 271
Db      721 GGAGAACCAACCACTGCTC-----CTCGAG----- 747
Qy      272 SerValValProValLeuGlnLysValAsnSerThrThrThrLysProValLeuArgThr 291
Db      748 -----GAAATTCACCTACTACCAACCCAGTGTCTGGAACT 780
Qy      292 ProSerProValHisProThrGlyThrSerGlnProGlnArgProGlnAspCysArgPro 311
Db      781 CCTCACCCTGTGCAACCTTACCGGAGCATCTCAGCCCAAGACCAAGAGATTGTGGGCC 840
Qy      312 ArgGlySerValLysGlyThrGlyLeuAspPheLeuGluAspProLysLeuCysTyrLeu 331
Db      841 CCGTGCTCAGTGAAGGGGACCGGATTGACTTC-----GACCCCAACTCTGTACTTGG 894
Qy      332 LeuAspGlyLeuLeuPheLysValLysLeuValLysLeuValLysLeuValLysLeu 351
Db      895 CTAGATGGAATCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 954
Qy      352 PheSerArgSerValGluThrAlaAlaAsnLeuGlnAspProAsnGlnLeuTyrAsnGlu 371
Db      955 TTACAGCAGAGAGCGAGAGCTGCTGCCAACCTGACGAGACCCCAACCACTCTTACATAG 1014
Qy      372 LeuAsnLeuGlyArgArgGluGluTyrAspValLeuGlnLysValLysValLysVal 391
Db      1015 CTCAATCTAGAGCGGAGAGAGAAATATGACGCTTGGAGAAAGAGGCGGCTCGGATCCA 1074
Qy      392 GluMetGlyGlyLysGlnGlnArgArgArgAsnProGlnGlnGluValTyrAsnAlaLeu 411
Db      1075 GAGATGGAGGCAACACACAGAGAGAGAACCCCGAGAGGCAATTCATCAATGCACTG 1134
Qy      412 GlnLysAspLysMetAlaGlnAlaTyrSerGlnLysGlnLysGlnLysGlnLysGln 431
Db      1135 CAGAAAGCAACATGGCGAAGCCTACAGTGAAGTCCGCAACAAAGCGAGCGGAGAG 1194
Qy      432 GlnLysGlnLysAspGlyLeuTyrGlnGlyLeuSerThrAlaThrLysAspThrTyrAsp 451
Db      1195 GCGAAGGGGCGAGTGGCTTTACCGAGGCTCTCAGCACTGCCACCAAGACACCTATGAT 1254
Qy      452 AlaLeuHisMetGlnThrLeuAlaProArg 461
Db      1255 GCCCTGCATATGACAGACCTGGCCCTCTGC 1284

```

RESULT 5
AAQ28262
ID AAQ28262 standard; DNA; 2011 BP.

XX AAQ28262;
AC
XX 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 11-FEB-1993 (first entry)
XX
XX Fv(FRP5) - ETA fusion gene.
XX
XX Monoclonal antibody; light chain; heavy chain; tumour; c-erbB-2;
KW variable region; ETA; ss.
XX
XX Pseudomonas aeruginosa; PAK.
OS
XX
FH Key Location/Qualifiers

```

FT sig_peptide 1..63
FT /tag= a
FT /label= ompA_signal_peptide
FT misc_RNA 64..87
FT /tag= b
FT /note= "FLAG peptide and enterokinase cleavage site"
FT misc_RNA 97..453
FT /tag= C
FT /label= FRP5_heavy_chain_variable_domain
FT misc_RNA 454..498
FT /tag= a
FT /label= Linker
FT misc_RNA 499..822
FT /tag= a
FT /label= FRP5_light_chain_variable_domain
FT CDS 826..1911
FT /tag= a
FT /label= EPA_1574-1747
FT 1912..2012
FT /tag= a
FT /label= EPA_3'_non_coding_region

EP502812-A1.
XX
XX 09-SEP-1992.
XX
XX 27-JAN-1992; 92EP-00810056.
XX
XX 05-FEB-1991; 91EP-00810079.
XX
XX (CIBA ) CIBA GEIGY AG.
XX
XX Wells WS, Hynes NE, Harwerth I, Groner B, Hardman N, Zwickl M;
XX WPI; 1992-302096/37.
XX
XX P-PSDB; AAR26982.
XX
XX Recombinant antibodies directed to growth factor receptor C-ERBB-2 - for
XX PT diagnosing and treating tumours expressing C-ERBB-2 e.g. breast or
XX PT ovarian tumours.
XX
XX PS Disclosure; Page 47-52; 67pp; English.
XX
XX The sequences given in AAQ28261-3 encode part of the exotoxin A (ETA)
XX CC sequence corresponding to positions 1574-1747 of the full exotoxin A
XX CC sequence. These sequences were used in the construction of Fv(FRP5)-ETA
XX CC fusion genes. The ETA sequence was used as a marker gene so that E. coli
XX CC transformed with the fusion gene could be identified. The fusion genes
XX CC were expressed in E. coli and the antibody was extracted. This
XX CC recombinant antibody can be used for the qualitative and quantitative
XX CC determination of c-erbB-2. This can be used for monitoring or in-vivo
XX CC localisation of tumours overexpressing c-erbB-2. (Updated on 25-MAR-2003
XX CC to correct PN field.) (Updated on 24-CT-2003 to standardise OS field)
XX
XX SQ Sequence 2011 BP; 383 A; 653 C; 620 G; 355 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3,85e-94 Length: 2011
Score: 1314.50 Matches: 281
Percent Similarity: 68.198 Conservative: 17
Best Local Similarity: 64.308 Mismatches: 50
Query Match: 53.854 Indels: 91
DB: 2 Gaps: 8

US-09-596-774-6 (1-461) x AAQ28262 (1-2011)
Qy 15 LysValProLysGlnLysLeuValGlnSerGlyProGlnLeuLysProGlnLysGln 34
Db 85 AAGCTAGCTTCTCAGAGTACCACTGACAGCAGTGTGACCTGAAGTGAAGAGCTGAGAG 144
Qy 35 ThrValLysLysSerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTrp 54
Db 145 ACGATCAAGATCTCTGCAAGGCTCTGGGTATCTTCTTCAACCACTATGAAATGAGCTGG 204

```

```

QY 55 ValIysGIAlaProGIyGInGIyLeuYsTrpMetGIyTrpIleAsnThsSerThrgly 74
DB 205 GTGAAGCAGGCTCCAGACAGAGGTTTAAAGTGAAGGCTGATTAACCTCCCTCAGCA 264
QY 75 GluSerThrPheAlaAspAspPheLysGIyArgPheAspPheSerLeuGIuThsSerAla 94
DB 265 GAGTCAACATTGCTGATGACTTCAAGGAGCGTTTGACTTCTTTGGAAACCTCTGCG 324
QY 95 AsnThrAlaTrpLeuGIuIleAsnAsnLeuLysSerGIuAspMetAlaThrTrpPheCys 114
DB 325 AACACTGCTATTTCAGATGACAACTCAAAAGTGAAGACATGCTACATATTCTGCT 384
QY 115 AlaArgTrpGIuValTrpPheGIyTrpValProGIyTrpGIyGInGIyThsThrValThr 134
DB 385 GCAGATGGAGGTTTACACAGGCTACGTTCTTCTGAGGACCAAGAGACAGGCTACAC 444
QY 135 ValSerSerGIyGIyGIySerGIyGIyGIySerGIyGIyGIySerAlaIle 154
DB 445 GTTTCCTCTGGCGGTGGCGGTCTGTGGCGGTGGCGGTGGCGGTGGCGGTCTGACATC 504
QY 155 GluLeuThrGlnSerHisLysPheLeuSerThrSerValGIyAspArgValSerIleThr 174
DB 505 CAGCTGACCCAGTCTCACAAATTCCTGCTCAGTACAGAGACAGGCTCAGCATCAC 564
QY 175 CysIlyAlaSerGIuAspValTrpAsnAlaValAlaTrpTrpGlnLysPheProGIyGIn 194
DB 565 TGCAAGGCGCAGCAGAGATGTATATGCTGTGCTGCTGATCAACAGAACAGAGACAA 624
QY 195 SerProLysLeuLeuIleTrpSerAlaSerSerArgTrpThrglyValProSerArgPhe 214
DB 625 TCTCTTAACCTTGATTTACTCGGCATCTCCGGTACACAGAGTCCCTCTGCTTC 684
QY 215 ThrGIySerGIySerGIyProAspPheThrPheThrIleSerSerValGIuAlaGIuAsp 234
DB 685 ACTGGCAGTGGCTCGGGCCGGATTTCATTTCACATCAGCAGTGTGAGGCTGAAAGAC 744
QY 235 LeuAlaValTrpPheCysGInGInHisPheArgTrpPheThrPheGIySerGIyThr 254
DB 745 CTGGCAGTTTATTCTGTGACAGAACATTTCTACTCCATTCAGTCTGGCGGAGACA 804
QY 255 LysLeuGIuIleLysAlaLeuGIuIleSerAsnSerValMetTrpPheSerSerVal 274
DB 805 AAATTGGAGATCAAGGCTTAGAG----- 828
QY 275 ProValLeuGIuLysValAsnSerThrThrTrpLysProValLeuArgThrProSerPro 294
DB 829 -----GGCGGACAGCTGCGCGCGCTGAC-CGCGCAGCAGGCGCTG 866
QY 295 ValHisProThrglyThrSerGIuProGIuArgProGIuAspCysArgProArgIySer 314
DB 867 CCACCTCCCGCTGGAGACTTTCACCCGTCATCGCA-----GCCGCG 908
QY 315 ValIysGIyThrglyLeuAspPheLeuGIuAspProLysLeuGIyTrpLeuAspGIy 334
DB 909 CGGCTGGGAACA-----ACTGGA 926
QY 335 IleLeuPheIleTrpGIyValIle-IleThrAlaLeuTrpIleuArgAlaLysPheSer-- 353
DB 927 GCAGTGGCGCTATCCGGTGCACCGGCTGTGCTCCCTCACTACGCGGCGGCTGTGATG 986
QY 354 -----ArgSerAlaGIuThrAla----- 359
DB 987 GAACCAAGTCGACGAGTGATCCGCAACCCCTGGCCGAGCCCGGCGAGCGCGGCACT 1046
QY 360 -----AlaAsnLeuGlnAspProAsnGIuLeuTrpAsnGIuLeuAsnLeuGIyArgArg 377
DB 1047 GGGCGAAGCATCCCGCAGACCGCGAGCAGCCGCTGCGCTGACCTT-GGCGCGCG 1105
QY 377 gGIuGIuTrpAspValLeuGIuLysValArgAlaArgAspProGIuMetGIyGIyLysGI 397
DB 1106 CCGAAGAG-----CGAGCGCTGCTGCTCG-----GCAGGCGCAGCG 1138

```

```

QY 397 ngIuArgArgAsnProGIuGIuGIyValTrpAsnAlaLeuGIuLysAspLysMetAl 417
DB 1139 GCAAGCAGCA----- 1148
QY 417 agIuAlaTrpSerGIuIleGIyThrLysGIyGluArgArgArgGIyLys 433
DB 1149 -----GGCCGCGCGCGGCAACGCGCAGCTGTGTAG 1178

RESULT 6
AAT42036
ID AAT42036 standard; DNA; 2070 BP.
AC AAT42036;
XX 29-JAN-1997 (first entry)
DT
DE Plasmid pSW202-5 fragment encoding scFv(FRP5)-ETA.
XX
XX Single chain antibody; scFv; monoclonal antibody; Mab; EGF; erbB-2;
KM epidermal growth factor; receptor; plasmid pSW202-5; cancer; therapy;
XX antitumour; exotoxin A; ETA; ss.
XX
XX Mus; SP.
OS Pseudomonas; aeruginosa.
OS Synthetic.
OS Chimeric.
XX
XX Key Location/Qualifiers
FH CDS 1..1956
FT 1..1956
FT sig_peptide 1..63
FT /tag= a
FT /note= "Omiga signal peptide"
FT mat_peptide 64..1953
FT /*tag= c
FT /product= "scFv(FRP5)-ETA"
XX
XX EP739984-A1.
XX
XX 30-OCT-1996.
PD
XX
XX 26-APR-1995; 95EP-00106275.
PF
XX
XX 26-APR-1995; 95EP-00106275.
PR
XX
XX (SANT-) SAN TUMORFORSCHUNGS GMBH.
PA
XX
XX Wels W, Schmidt M, Groner B;
XX
XX MPI; 1996-478748/48.
DR P-PSDB; AAM05156.
XX
XX Bivalent fusion proteins that bind epidermal growth factor receptor or
PT analogues - and comprise at least two different cell surface binding
PT domain(s), useful for tumour therapy.
XX
XX Example 7; Page 21-22; 52pp; English.
PS
XX
XX A fragment (AAT42036) of plasmid pSW202-5 codes for scFv(FRP5)-ETA
CC (AAM05136), comprising the single-chain binding region of murine
CC monoclonal antibody FRP5, which is specific for human epidermal growth
CC factor receptor erbB-2, joined to exotoxin A (ETA). A fragment of plasmid
CC pMW52-5 (see also AAT42036) encoding scFv(FRP5) (AAM05134) was ligated
CC to pSW202-5, a plasmid contg. the Pseudomonas aeruginosa PAK ETA gene, to
CC produce pSW202-5. This can be utilised in the construction of bacterial
CC expression vectors (see also AAT42037-44) encoding bivalent fusion
CC proteins (AAM05137-44) useful as antitumour agents
XX
SQ Sequence 2070 BP; 402 A; 675 C; 628 G; 365 T; 0 U; 0 Other;

Alignment Scores: 5.74e-94 Length: 2070
Pred. No.: 1312.50 Matches: 284

```

Percent Similarity: 68.59% Conservative: 13
 Best Local Similarity: 65.59% Mismatches: 55
 Query Match: 53.77% Indels: 83
 DB: 2 Gaps: 8

US-09-596-774-6 (1-461) x AAT42036 (1-2070)

```

Oy 19 GlnIleGlnLeuValGlnSerGlyProGluLeuLysLysProGlyGluThrValLysIle 38
Db 115 CAGGTACCACTCCAGACGCTGAGCCTGAACTGAAAGACCTGGAGAGCACTCAAGATC 174
Oy 39 SerGlyLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValLysGlnAla 58
Db 175 TCCTCCAGAGCCTCTGGGTATCCTTCAACAATGAAATGAACTGGGTGAAGACGCT 234
Oy 59 ProGlyGlnGlyLeuLysTyrMetGlyTyrPheLeuThrSerThrGlyLysThrPhe 78
Db 235 CCAGGACAGGGTTTAAAGTGAAGGCTGGATTAAACCTCCACTGGAGAGTCAACATTT 294
Oy 79 AlaAspAspPheLysGlyLysPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 98
Db 295 GCTGATGACTTCAAGGAGCGTTTGACTTCTTGGAAACCTTGCCAAACATCTCCTAT 354
Oy 99 LeuGlnIleAsnAsnLeuLysSerGlyLysPheAlaThrTyrPheCysAlaArgTyrGlu 118
Db 355 TTGCAGATCAACAACTCAAAAGTGAAGACATGCTCACTATTTCTGCAAGATGGAG 414
Oy 119 ValTyrHisGlyTyrValProTyrTyrGlyGlnGlyThrThrValThrValSerSerGly 138
Db 415 GTTTTCCACGCGCTTACGCTTCTTACTGGGGCCAAAGGACCGCTACCGTTCTCTGGC 474
Oy 139 GlyGlyGlySerGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThrGln 158
Db 475 GGTGGCGGTTCTGGTGGCGGTGCTCCGCGGTGGCGGTTCGATCGACTGACCGAG 534
Oy 159 SerHisLysPheLeuSerThrSerValGlyAspAlaGlySerIleThrCysLysAlaSer 178
Db 535 TCTCACAATTCCTGCTCCTCACTCACTAGAGACAGGGTCACTCACTCGCAAGCCAGT 594
Oy 179 GlnAspValTyrAsnAlaValAlaTyrTyrGlnGlnLysProGlyGlnSerProLysLeu 198
Db 595 CAGAGTGTATATGCTGCTGCTCCGATATCAAGAAACCAAGCAATCTCTTAACCTT 654
Oy 199 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 218
Db 655 CTGATTACTGGGATCTCTCCCGGTACACTGAGATCCCTTCTCGCTTCACTGGCAGTGGC 714
Oy 219 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 238
Db 715 TCTGGGCGGATTCACCTTCACTCACTCACTGAGCTGCGTGAAGACTGGCAGTTTAT 774
Oy 239 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyTyrThrLysLeuGluIle 258
Db 775 TTCTGTACGCAACATTTTCTGATCTCATTCACGTTCCGCTGGGGAACAATTTGGAATC 834
Oy 259 LysAlaLeuGluIleSerAsnSerValMetTyrPheSerSerValProValLeuGln 278
Db 835 AAAGCTTAGAGCA----- 848
Oy 279 LysValAsnSerThrThrThrLysProValLeuArgThrProSerProValHisProThr 298
Db 849 -----CCATATCAACATCACTCAAGA 869
Oy 299 GlyThrSerGlnProGlnArgProGluAspCysArgProArgGlySerValLysGlyThr 318
Db 870 GGG-----CGGACGCTGGCGCGCTGAC-----CGGACCAACAGCCCTGCCACCTGCCCT 920
Oy 319 GlyLeuAspPheLeuGluAspProLysLeuGlyTyrLeuLeuAspGlyIleLeuPheIle 338
Db 921 GGAGACTTTTCACTCCGCTATCGCGCGCGCTGGGAACAACCTGAGAGAGATGCGGCTA 980
Oy 339 TyrGlyValIle-IleThrAlaLeuTyrLeuArgAlaLysPheSer----- 353
  
```

```

Db 981 TCCGATGACGGCGGTGCTGCTTCACTGACCGCGGCACTGTCATGAAACAGTCA 1040
Oy 354 -----ArgSerAlaGluThrAla-----AlaAs 361
Db 1041 CCAGGTGATCCGCAACGCTCGGCGGACCGCGAGCGGCGAGCACTGGGCGAAGCGAT 1100
Oy 361 nLeuGlnAspProAsnGlnLeuTyrAsnAlaLeuAsnLeuGlyLysArgGlyGluTyrAs 381
Db 1101 CCGGAGAGCGCGGAGCGAGCGGCGCTGCGCTGACCTT-GGCGCGCGCGAGG----- 1154
Oy 381 pValLeuGluLysLysArgAlaArgAspProGluMetGlyGlyLysGlnGlnArgArg 401
Db 1155 -----CGAGGCTTGTCTCG-----CGAGGAGACCGGCAACGAGCA-- 1190
Oy 401 gAsnProGlnGluGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaGluAlaTyrSe 421
Db 1190 ----- 1190
Oy 421 rGluIleGlyThrLysGlyGluArgArgArgLys 433
Db 1191 -----GGCGGCGCGGCGCAAGCGAGCGTGTGAG 1220

RESULT 7
AAT42044
ID AAT42044 standard; DNA; 2730 BP.
XX
AC AAT42044;
XX
XX 29-JAN-1997 (first entry)
XX
DE Plasmid pMS246-5-5 encoding scFv2(FRP5)-ETA (version 3).
XX
KW Single chain antibody; scFv; monoclonal antibody; Mb; EGF; erbB-2;
KW epidermal growth factor; receptor; plasmid pMS246-5-5; cancer;
KW exotoxin A; ETA; antitumour; ss.
XX
OS Mus; sp.
OS Pseudomonas; aeruginosa.
OS Synthetic.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT CDS 1..2700
FT sig_peptide 1..63
FT mat_peptide /note= "OmpA signal peptide"
FT /tag= c
FT /product= "scFv2 (FRP5/FRP5)-ETA"
XX
XX EP739984-A1.
XX
PD 30-OCT-1996.
XX
PF 26-APR-1995; 95EP-00106275.
XX
PR 26-APR-1995; 95EP-00106275.
XX
PA (SANT-) SAN TUMORFORSCHUNGS GMBH.
XX
PI Wels W, Schmidt M, Groner B;
XX
DR WPI: 1996-478748/48.
XX
DB P-PSDB; AAM05144.
XX
XX Bivalent fusion proteins that bind epidermal growth factor receptor or
XX analogues - and comprise at least two different cell surface binding
XX domain(s), useful for tumour therapy.
XX
PS Example 12; Page 43-45; 52pp; English.
XX
CC A fragment (AAT42044) of bacterial expression plasmid pMS246-5-5 codes
  
```

for scfv2(FRPs/225)-ETA (AAW05144), and is obt'd. by ligating 2 copies of
 CC DNA coding for the single-chain binding region of murine monoclonal
 CC antibody FRP5 (see also AAT2034) and portions of *Pseudomonas aeruginosa*
 CC exotoxin A. pMS246-5-5 can be utilised in the prodn. of bivalent fusion
 CC protein in bacterial (esp. *E. coli*) host cells. Such fusion proteins (see
 CC also AAW05138-43) bind the epidermal growth factor receptor and are
 CC useful as antitumour agents

Sequence 2730 BP; 567 A; 832 C; 798 G; 533 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	8.02e-94	Length:	2730
Score:	1312.50	Matches:	284
Percent Similarity:	68.59%	Conservative:	13
Best Local Similarity:	65.59%	Mismatches:	55
Query Match:	53.77%	Indels:	83
DB:	2	Gaps:	8

US-09-596-774-6 (1-461) x AAT2044 (1-2730)

```

OY 19 GlnIleGlnLeuValGlnSerGlyProGluLeuValysProGlyValThrValylsile 38
DB 115 CAGGTACCACTGCACAGCTGACCTGAACTGAAGAACCTGGAGAGCACTCAAGATC 174
OY 39 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValylsGlnAla 58
DB 175 TCCTCCAAAGGCTCTGGGTATCTTTCACAACTATGAACTATGAACTGGGTGAAGCAGCT 234
OY 59 ProGlyGlnGlyLeuValTyrPheMetGlyTyrPheAsnThrSerThrGlyLysSerThrPhe 78
DB 235 CCAGACACGGGCTTAAAGTGAAGTGGCTGATTAACCTCCACCTGAGAGTCAACATTT 294
OY 79 AlaAspAspPheLysGlyLysArgPheAspPheSerLeuGluThrSerAlaAsnThrLysTyr 98
DB 295 GCTGATGACTTCAAGAGGAGCGTTTACCTTCTTGGAAACCTCTGCAACACTGCTAT 354
OY 99 LeuGlnIleAsnAsnLeuLysSerGlyLysAspMetAlaThrTyrPheCysAlaArgTyrGlu 118
DB 355 TTGCAGATCAACAACCTCAAAAGTGAAGCAGGCTACATATTTCTGTGCAAGATGGAG 414
OY 119 ValTyrHisGlyTyrValProTyrTyrPheGlyGlnGlyLysThrValThrValSerGly 138
DB 415 GTTACACAGGCTACGTTCTTACCTGAGGCGCAAGGAGCACCGTCCCTCTGCGC 474
OY 139 GlyGlyGlySerGlyGlyGlySerGlyGlyGlySerGlyGlySerProGlyGlnLeuThrGln 158
DB 475 GGTGGCGGTTCTGGTGGCGGTGCTCCGCGGTGGCGGTCTGACATCCAGCTGACCGAG 534
OY 159 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSer 178
DB 535 TCTCAAAATTCCTCTCCACTTCAAGAGAGCAGGCTCAGATCACTCGCAAGGCCACT 594
OY 179 GlnAspValTyrAsnAlaValAlaTyrPheGlyGlnGlyLysProGlyGlnSerProLysLeu 198
DB 595 CAGGATGTGTAATGCTGTGCTGCTGATCAACAGAAACAGCAATCTCTAAACTT 654
OY 199 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 218
DB 655 CTGATTACTCGGCATCTCTCCCGGTACACTGAGAGTCCCTTCTGCTTCACTGGCAGTGGC 714
OY 219 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 238
DB 715 TCTGGGCGGATTTCACTTTCACATCAGCATGAGTGGAGGCTGAAAGACTGGCAGTTTAT 774
OY 239 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyTyrLysLeuGluIle 258
DB 775 TTCTGTCAAGCAACATTTCTGACTCATTCAGTTCCGCTCGGGGACAAATTTGAGATG 834
OY 259 LysAlaLeuGluLysSerAsnSerValMetTyrPheSerSerValValProValLeuGln 278
DB 835 AAAGCTCTAGAGCA----- 848
OY 279 LysValAsnSerThrThrLysProValLeuArgThrProSerProValHisProThr 298

```

```

DB 849 -----|||||
OY 229 GlyThrSerGlnProGlnArgProGluAspCysArgProArgGlySerValylsGlyThr 318
DB 870 GGG---CGGCGGCTTGGCGCGGTAC-----CGCGACCAAGGCTGCCACTGCCCT 920
OY 319 GlyLeuAspPheLeuGluAspProLysLeuCyTyrLeuLeuAspGlyIleLeuPheIle 338
DB 921 GGAGACTTTCACCCCTCATCCGACCCCGCGGCTGGGAAACACTGAGACACTGCCGCTA 980
OY 339 TyrGlyValIle-IleThrAlaLeuTyrLeuArgAlaLysPheSer----- 353
DB 981 TCCGGTGCAGGGCTGTGCTGCTTACTGCGCGGCGGACTGATGAAACAGATCGA 1040
OY 354 -----ArgSerAlaGluThrAla-----AlaAs 361
DB 1041 CCAGGTGATCCGCAACGCCCTGGCCAGCCCGGCGAGCGGCGACTGGGCGAAGCGAT 1100
OY 361 nLeuGlnAspProAsnGlnLeuTyrAsnGluLeuAsnLeuGlyArgArgGluTyrAs 381
DB 1101 CCGGAGCAGCCCGAGAGCGGCCCTGCTGCTGACCTT-GGCCCGCGCGAGAG----- 1154
OY 381 pValLeuGluLysLysArgAlaArgAspProGluMetGlyLysGlnGlnArgArg 401
DB 1155 -----CGAGGCTTCTGCTCG-----GCAGGCGACCGGCAACGAGCA-- 1190
OY 401 GAsnProGlnGluValTyrAsnAlaLeuGlnLysAspLysMetAlaGluAlaTyrSe 421
DB 1190 ----- 1190
OY 421 rGluIleGlyThrLysGlyGluArgArgGlyLys 433
DB 1191 -----GGCGGCGCGGCAACGCCGAGCTGTGAG 1220

```

RESULT 8
 AAT29413
 ID AAT29413 standard; DNA; 1862 BP.
 XX
 AC AAT29413;
 XX
 DT 16-OCT-2003 (revised)
 DT 19-AUG-1996 (first entry)
 XX
 DE DGA4-DT-scfv(FRP5) gene in pSM5-GDS.
 XX
 KW Nucleic acid transfer system; gene transfer; gene therapy;
 KW cell targeting; multidomain protein; vector; cancer; GAL4;
 KW diptheria toxin; single chain antibody; scfv; FRP5; ss.
 XX
 OS Saccharomycetes; cerevisiae.
 OS Corynebacterium; diptheriae.
 OS Mus sp.
 OS Chimeric.
 OS Synthetic.

Key	Location/Qualifiers
FT CDS	1..1908
FT FT	/*tag= a
FT FT	/product= "GAL4-DT-scfv(FRP5) fusion protein"
XX	MO9613599-AI.
XX	09-MAY-1996.
XX	31-OCT-1995; 95WO-EP004270.
XX	01-NOV-1994; 94EP-00810627.
XX	(WELLS/) WELLS W.
XX	Wells W, Fominaya J;
XX	


```

QY 39 SerCysAlaSerGlyThrProPheThrAsnTyrGlyMetAsnTrpValLysGlnAla 58
DB 175 TCTGTGAAGCCCTCTGGTATCCTTTCAACAATGATGATGACTGGGTGAAGCAGGCTT 234
QY 59 ProGlyGlyGlyLeuLysTrpMetGlyTrpLLeaSerThrGlyLysSerThrPhe 78
DB 235 CCAGGACAGGCTTTAAAGTGAATGGGCTGATTTAAACATCCACATGAGAGTCAACATTT 294
QY 79 AlaAspAspPheLysGlyThrPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 98
DB 295 GCTGATGACTTCAAGGAGCAGGTTTGACTTCTTTGGAAACCTCTGCCAAACCTCCAT 354
QY 99 LeuGlnLLeaAsnLeuLysSerGlyAspMetAlaThrTyrPheCysAlaArgTrpGlu 118
DB 355 TTGCAGATCAACACCTCAAAAGTGAAGCATGGCTACATATTTCTGTCCAGAGAGGAG 414
QY 119 ValTyrHisGlyTyrValProTyrTrpGlyGlnGlyThrValThrValSerSergly 138
DB 415 GTTTACACAGGCTACGTTCTTACTGGGCGCAAGGAGCACAGCTACCGTTCTCTGGC 474
QY 139 GlyGlyGlySerglyGlyGlySerglyGlyGlySerglySerglyGlnLeuThrGln 158
DB 475 GGTGGCGGTTCTGTGGCGGCTGCTCCGCGGAGGCGGTTCTGACATCCAGCTGACCAG 534
QY 159 SerHisLysPheLeuSerThrSeryAlaGlyAspArgValSerLLeaThrCysLysAlaSer 178
DB 535 TCTCACAATTCCTGCTCCACCTTCAGTAGAGACAGGGTCAGATCACTGCAAGCCAGT 594
QY 179 GlnAspValTyrAsnAlaValAlaTrpTyrGlnGlnLysProGlyGlnSerProLysLeu 198
DB 595 CAGGATGATGAATGAATGCTGTGCTGCTGATCAACAGAAACCAAGCAATCTCCATAACTT 654
QY 199 LeuLLeaTyrSerAlaSerSerglyTyrThrGlyValProSeryPheThrGlySergly 218
DB 655 CTGATTTACTGGGACTCTCCCGGTAACACTGAGTCCCTTCTCGCTTCACTGGCAGTGGC 714
QY 219 SerGlyProAspPheThrPheThrLLeaSerValGlnAlaGlnAspLeuAlaValTyr 238
DB 715 TCTGGGCGGATTCACCTTCACCATCAGAGTGGCAGGCTGAAAGACTGCGAGTTTAT 774
QY 239 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerglyThrLysLeuGlnLLea 258
DB 775 TTCTGTCAAGACATTTCTGATCTCATTCACGTTCTGGGAGCAAAATTTGAGATCTC 834
QY 259 LysAlaLeuGlnLLeaSerAsnSerValMetTyrPheSerSeryValProValLeuGln 278
DB 835 AAAGCTTAGAGCA----- 848
QY 279 LysValaAsnSerThrThrLysPheProValLeuArgThrProSeryProValHisProThr 298
DB 849 -----CCATCATCACTCACTCAGTAGA 869
QY 299 GlyThrSergLInProGlnArgProGlnAspCysArgProArgLysSeryValLysGlyThr 318
DB 870 GGG---CGGCAAGCTGGCGCGCGTGAC-----CGCGACCAAGGCTGCGCACCTGCCGCT 920
QY 319 GlyLeuAspPheLeuGlnAspProLysLeuCysTyrLLeuAspGlyLLeaPheLLea 338
DB 921 GGAGACTTTCAACCGTCACTCGCAGCGCGCGGCTGGGAAACAATGGAGCAGTCCGCTA 980
QY 339 TyrGlyValLLeaLLeaThrAlaLeuTyrLLeaArgAlaLysPheSeryArgSeryAlaGln 358
DB 981 TCCGCTGACAGCGGCTGTGCTGCTCTCACTGCGCGCGCGAGTGTCA----- 1026
QY 358 TAlaAlaAsnLeuGlnAspProAsnGlnLeuTyrAsnGlnLeuAsnLeuGlyArgArgGly 378
DB 1027 -----TGAAACAGGTGACCA-GGT----- 1046
QY 378 uGluTyrAspValLeuGlnLysLysArgAlaArgAspProGlnMetGlyLysGlnGly 398
DB 1047 -----GATCCGCAAGCGCTGCGCGAGCCCG 1072
QY 398 nArgArgArgAsnPro-GlnGlnGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaG 418

```

```

DB 1073 GCAAGCGGCGGAGCTTGGCGGAGCGATCCGCGGACAGCCGCGAGCCCGCTGTGGCC 1132
QY 418 lualaTyrSergLInLLeaGlyThrLysGlyGlnAspArgArgGlyLysGlyHisAsp 436
DB 1133 TGACCCCTGCGCGCGCGGAGAGCAGAGCGCTTCGTCGCGGAGGCGACCGGACAGCAG 1168

RESULT 11
AAT42039
ID AAT42039 standard; DNA; 2793 BP.
XX
AC AAT42039;
XX
DT 29-JAN-1997 (first entry)
XX
DE Plasmid pMS238-5-225 fragment encoding scFv2 (FRP5/225)-ETA.
XX
KW Single chain antibody; scFv; monoclonal antibody; Mab; EGF; erbB-2;
KW epidermal growth factor; receptor; plasmid pMS238-5-225; cancer;
KW exotoxin A; ETA; antitumour; ss.
XX
OS Mus; sp.
OS Pseudomonas; aeruginosa.
OS Synthetic.
OS Chimeric.
XX
FH Key
FT CDS
FT 1..2679 Location/Qualifiers
FT /*tag= a
FT sig_peptide 1..63
FT /*tag= b
FT /note= "Ompa signal peptide"
FT mat_peptide 64..2676
FT /*tag= c
FT /product= "scFv2 (FRP5/225)-ETA"

EP739984-A1.
XX
PD 30-OCT-1996.
XX
PF 26-APR-1995; 95EP-00106275.
XX
PR 26-APR-1995; 95EP-00106275.
XX
PA (SANT-) SAN TUMORFORSCHUNGS GMBH.
XX
PI Wels W, Schmidt M, Groner B;
XX
DR WPI; 1996-478748/48.
XX
DR P-PSDB; AAM05139.
XX
PT Bivalent fusion proteins that bind epidermal growth factor receptor or
PT analogues - and comprise at least two different cell surface binding
PT domain(s), useful for tumour therapy.
XX
PS Example 11; Page 28-30; 52pp; English.
XX
CC A fragment (AAT42039) of bacterial expression plasmid pMS238-5-225 codes
CC for scFv2 (FRP5/225)-ETA (AAM05139), and is obt'd. by ligating DNA
CC fragments coding for the single-chain binding region of murine monoclonal
CC antibody FRP5 (see also AAT42034), portions of Pseudomonas aeruginosa
CC exotoxin A and the single-chain binding region of murine monoclonal
CC antibody 225 (see also AAT42033). pMS238-5-225 can be utilised in the
CC prodn. of bivalent fusion protein in bacterial (esp. E. coli) host cells.
CC Such fusion proteins (see also AAM05138-44) bind the epidermal growth
CC factor receptor and are useful as antitumour agents
XX
SQ Sequence 2793 BP; 593 A; 847 C; 806 G; 547 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1,87e-93 Length: 2793
Score: 1308.00 Matches: 275
Percent Similarity: 71.19% Conservative: 24

```


CC also AAW05138-44) bind the epidermal growth factor receptor and are
 CC useful as antitumour agents
 XX
 SQ Sequence 2793 BP; 578 A; 858 C; 812 G; 545 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	Length:	Score:	Matches:
1	87e-93	1308.00	275
Percent Similarity:	71.19%	Conservative:	24
Best Local Similarity:	65.48%	Mismatches:	59
Query Match:	53.58%	Indels:	64
DB:	2	Gaps:	5

US-09-596-774-6 (1-461) x AAT42043 (1-2793)

```

OY 19 GlnIleGlnLeuValGlnSerGlyProGluLeuIleuValysProGlyValThrValysIle 38
DB 115 CAGGTACAACTGCAGACGCTGACCTGAACCTGAAGAACTCGAGAGACAGTCAAGATC 174
OY 39 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValIleGlnAla 58
DB 175 TCCGTCAAGGCGCTCGGGATCTTTCACAACTGATGATCACTGGGTGAAGAGGCT 234
OY 59 ProGlyGlnGlyLeuIleTyrMetGlyTyrIleAsnThrSerThrGlyIleSerThrPhe 78
DB 235 CCAGACAGGCGTTTAAAGTGAAGTGGCTGATTAACACCTCCAGCTGAGAGTCAACATTT 294
OY 79 AlaAspAspPheLysGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 98
DB 295 GCTGATGACTTCAAGAGGAGCGGTTGACTTCTTTGGAACCTCGCAACACCTGCTAT 354
OY 99 LeuGlnIleAsnAsnLeuIleSerGlyAspMetAlaThrTyrPheCysAlaArgTyrGlu 118
DB 355 TTGCAGATCAACAACCTCAAAAGTGAAGACATGGCTATATTTCTGTGCAAGATGGAG 414
OY 119 ValTyrHisGlyTyrValProTyrTyrGlyGlnGlyThrValThrValSerSerGly 138
DB 415 GTTACCAACGCGCTACGTTCTTACTGGGGCCAAAGGACACAGTCAACCGTTCTCTGGC 474
OY 139 GlyGlyGlySerGlyGlyGlySerGlyGlyGlyGlySerAspIleGlnLeuThrGln 158
DB 475 GGTGGCGGTTCTGTGGCGGAGGCTCCGCGCGGTGCTTGCACATCCAGTCAACCCAG 534
OY 159 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysAlaSer 178
DB 535 TCTCAACAATTCCTGTCCACTTCAGTAGGAGACAGGGGTCAACATCCTGCAAGCCAGT 594
OY 179 GlnAspValTyrAsnAlaValAlaTyrTyrGlnGlnLysProGlyGlnSerProLysLeu 198
DB 595 CAGAGTGTGATTAAGCTGTGCTGTGATCAACAGAAACAGAGACATCTCCCTAAACTT 654
OY 199 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 218
DB 655 CTGATTAATCTGGCATCTCCCGGTACACTGGATCTCTCTCCCTTCACTGGCAGTGGC 714
OY 219 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 238
DB 715 TCTGGCGCGGATTTCACTTTCACATCAAGCAGTGTGAGGCTGAAGACTGGGCGATTAT 774
OY 239 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyTyrThrLysIleGlnIle 258
DB 775 TTCTGTGACGACAACTTTCGACTCCATCACTGCTCGGCGGACAAATGAGATC 834
OY 259 LysAlaLeuGlnIleSerAsnSerValMetTyrPheSerSerValProValLeuGln 278
DB 835 AAGCTCTAGAGCA----- 848
OY 279 LysValAsnSerThrThrThrLysProValLeuArgThrProSerProValHisProThr 298
DB 849 -----CCATCATCACCATCACCAGTA 869
OY 299 GlyThrSerGlnProGlnArgProGlnAspCysArgProArgGlySerValLysGlyThr 318
  
```

```

DB 870 GGG---CGGACGCTGGCGCGCTGAC-----CGCGACCAAGGCTGCCACTGCCGCT 920
OY 319 GlyLeuAspPheLeuGluAspProLysAlaCysTyrLeuLeuAspGlyIleLeuPheIle 338
DB 921 GGAGACTTTCAACCCGTCATCGCCAGCGCGGTGGGAACAACATGGAGACAGTGGGCTA 980
OY 339 TyrGlyValIle-IleThrAlaLeuTyrIleuArgAlaLysPheSerArgSerAlaGluThr 358
DB 981 TCCGATGACGCGGCTGTGGCTCTTACTTGTGGCGCGCGACTGCA----- 1026
OY 358 rAlaAlaAsnLeuGlnAspProAsnGlnLeuTyrAsnGluLeuAsnLeuGlyArgArgI 378
DB 1027 -----TGAATCAGGTGACCA-GGT----- 1046
OY 378 uGluTyrAspValLeuGlnLysValArgAlaArgAspProGluMetGlyLysGlnI 398
DB 1047 -----GATCCCAACGCGCTGCGCAGCCCG 1072
OY 398 nArgArgArgAspPro-GlnGluGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaG 418
DB 1073 GCAGCGCGCGGACTGTGGCGAAAGCATCGCGAGCAGCGCGAGAGAGCGCGCTGTGGCC 1132
OY 418 IuaLarYSerGluIleGlyThrLysGlyGluArgArgGlyLysGlyHisAsp 436
DB 1133 TGACCTGCGCGCGCGAGAGAGCGAGCGCTTGTGCGGAGGCGACCGGCAAGCAG 1188
  
```

RESULT 13

AAT42041
 ID AAT42041 standard; DNA; 3177 BP.

AC AAT42041;

DT 29-JAN-1997 (first entry)

XX Plasmid pMS240-5-225 fragment encoding scFv2 (FRP5/225)-ETA.

XX Single chain antibody; scFv; monoclonal antibody; Mab; EGF; erbB-2;

XX epidermal growth factor; receptor; plasmid pMS240-5-225; cancer;

XX exotoxin A; ETA; antitumour; ss.

XX Mus; sp.

XX Pseudomonas; aeruginosa.

XX Synthetic.

XX Chimeric.

XX Key

FT CDS

FT sig_peptide

FT mat_peptide

FT /tag= C

FT /product= "scFv2 (FRP5/225)-ETA"

XX EP39984-A1.

XX 30-OCT-1996.

XX 26-APR-1995; 95EP-00106275.

XX 26-APR-1995; 95EP-00106275.

XX (SANT-) SAN TUMORFORSHINDS GMBH.

XX wels w, Schmidt M, Groner B;

XX MPI; 1996-478748/48.

XX P-PSDB; AAW05141.

XX Bivalent fusion proteins that bind epidermal growth factor receptor or

XX analogues - and comprise at least two different cell surface binding

XX domain(s), useful for tumour therapy.


```

FT      sig_peptide      /tag= a
FT      23..85
FT      /*tag= b
FT      /label= ompa_signal_peptide
FT      89..445
FT      m1ec_RNA
FT      /*tag= c
FT      /label= FRP5_heavy_chain_variable_domain
FT      446..490
FT      /*tag= d
FT      /label= Linker
FT      491..814
FT      m1ec_RNA
FT      /*tag= e
FT      /label= FRP5_light_chain_variable_domain
FT      815..2155
FT      mat_peptide
FT      /*tag= f
FT      /label= phoa_coding_region

XX      EP502812-A1.
XX      09-SEP-1992.
XX      27-JAN-1992; 92EP-00810056.
XX      05-FEB-1991; 91EP-00810079.
XX      (CIBA ) CIBA GEIGY AG.
XX      Weis WS, Hynes NE, Harwerth I, Groner B, Hardman N, Zwickl M;
XX      WPI; 1992-302096/37.
XX      DR P-PSDB; AAR26980.
XX      Recombinant antibodies directed to growth factor receptor C-erbB-2 - for
XX      PT diagnosing and treating tumours expressing C-erbB-2 e.g. breast or
XX      PT ovarian tumours.
XX      PS Disclosure; Page 34-40; 67pp; English.
XX      CC The sequence given encodes the single chain recombinant antibody
XX      CC designated Fv(FRP5)-phoa. The alkaline phosphatase gene (phoa) was used
XX      CC as a marker gene so that E. coli transformed with the fusion gene could
XX      CC be identified. The fusion gene was expressed in E. coli and the antibody
XX      CC was extracted. This recombinant antibody can be used for the qualitative
XX      CC and quantitative determination of c-erbB-2. This can be used for
XX      CC monitoring or in-vivo localisation of tumours overexpressing c-erbB-2.
XX      CC (Updated on 25-MAR-2003 to correct FN field.)
XX      SQ Sequence 2233 BP; 566 A; 573 C; 606 G; 488 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3.54e-93 Length: 2233
Score: 1303.00 Matches: 249
Percent Similarity: 91.37% Conservative: 5
Best Local Similarity: 89.57% Mismatches: 9
Query Match: 53.38% Indels: 16
DB: 2 Gaps: 2

US-09-596-774-6 (1-461) x AAQ28257 (1-2233)
QY 1 MetAlaTPValTTPThLeuLeuPheLeuMetAlaAlaAlaValProLysGlnIle 20
DB 45 TTGCAGTGGCACTGGCTGTTCCGTACC-GTAGCGCAAGCT-----TCTCAGGTA 94
QY 21 GlnLeuValGlnSerGlyProGlnLeuLysProGlyGlnThrValLysIleSerCys 40
DB 95 CAACCTGCAGCACTGCAGCTGAAGAAAGCTGAGAGACAGTCAGATCTCTGCG 154
QY 41 LysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTPValLysGlnAlaProGly 60
DB 155 AAGGCTCTGGGTATCTTTCCAAACATATGAAATGAACTGGGTGAAGCAGGCTCCAGGA 214
QY 61 GlnGlyLeuLysTPMetGlyTTPLeuAsnThrSerThrGlyLysSerThrPheAlaAsp 80

```

```

DB 215 CAGGGTTTAAAGTGAATGGCTGGATTAAACCTCCAGCTGAGAGTCAACATTGGCTGAT 274
QY 81 AsPheLysGlyArgPheAspPheSerLeuGlnThrSerAlaAsnThrAlaTyrLeuGln 100
DB 275 GACTTCAGGAGACGGTTTGACTTCTTTGGAAACCTTGCAACACTGCTTATTGGAG 334
QY 101 IleAsnLeuLysSerGlyAspMetAlaThrTyrPheCysAlaArgTyrGlnValTyr 120
DB 335 ATCAACAACCTCAAAAGTAAGAACATGGCTCAATATTTCTGGCAAGATGGAGGTTTAC 394
QY 121 HisGlyTyrValProTyrTTPGlyGlnGlyThrThrValThrValSerSerGlyGlyGly 140
DB 395 CACGGCTACGTTCTTCTTACTGGGGCCAAAGGAGCAACGGTCAACCGTTTCTCTGGCGGTGCG 454
QY 141 GlySerGlyGlyGlySerGlyGlyGlyGlySerAspIleGlnLeuThrGlnSerHis 160
DB 455 GGTCTGTGGCGGGGTGGCTGGCGGGTGGCGGTTCTGACATTCAGTGAAGTCAAGTCTAC 514
QY 161 LysPheLeuSerThrSerValGlyAspArgValSerIleThrCysAlaSerGlnAsp 180
DB 515 AAATTCTGTCCTCACTTCAGTAGAGACAGGCTCAGCATCACTGCAGAGCCAGTCAGGAT 574
QY 181 ValTyrAsnAlaValAlaTyrTyrGlnGlnLysProGlyGlnSerProLysLeuLeuIle 200
DB 575 GTGTATAATGCTGTGCTGGTATCAACAGAAACAGAGCAATCTCTAAACTTCTGATT 634
QY 201 TyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGlySerGly 220
DB 635 TACTCGGCATCTCCCGGTACACTGGAAGTCCCTTCTGCTTCACTGGGAGGCTCTGGG 694
QY 221 ProAspPheThrPheThrIleSerSerValGlnAlaGlnAspLeuAlaValTyrPheCys 240
DB 695 CCGATTTCATCTTCACCATATGACAGTGTGCAAGCTGAAGACTGGCAGTTATTTCGT 754
QY 241 GlnGlnHisPheArgThrProPheThrPheGlySerGlyThrLysLeuGlnIleLysAla 260
DB 755 CAGCAACATTTTCGTACTCCATTCACGTTCCGCTCGGGGAGCAAAATTGGAGATCAAACT 814
QY 261 LenglIleSerAsnSerValMetCtyrPheSerSerValValProValLengIn 278
DB 815 CTAGAG-----CCGTGTTCTGGAA 832

```

Search completed: April 20, 2005, 10:44:24
 Job time : 767.356 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 20, 2005, 10:19:55 ; Search time 788.902 Seconds

(without alignments)
3545.124 Million cell updates/sec

Title: US-09-596-774-6

Perfect score: 2441

Sequence: 1 MAWWTLLFLMAAKVPRKQI.....ISTAKDTYDALHMQTLAPR 461

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 5622541 seqs, 303335566 residues

Total number of hits satisfying chosen parameters: 11245082

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODE=frame_plus_p2n.model -DEV=xlh
-O=/cgnt2_1/USPTO_epool_h/US05956774/runatc_20042005_075444_12560/app_query.fasta_1.1230
-DB=Published Applications NA -GFM=faetap -SUFFIX=apr20.rnpb -MINMATCH=0.1
-LOOPL=0 -LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=spc -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=plc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US05956774 -CGCN_1_1_598 -ORUNATC_20042005_075444_12560
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:*

- 1: /cgnt2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgnt2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgnt2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
- 4: /cgnt2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
- 5: /cgnt2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
- 6: /cgnt2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
- 7: /cgnt2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
- 8: /cgnt2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
- 9: /cgnt2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgnt2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgnt2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgnt2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
- 13: /cgnt2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
- 14: /cgnt2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
- 15: /cgnt2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
- 16: /cgnt2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
- 17: /cgnt2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
- 18: /cgnt2_6/ptodata/1/pubpna/US10F_NEW_PUB.seq:*
- 19: /cgnt2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
- 20: /cgnt2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
- 21: /cgnt2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
- 22: /cgnt2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1249	51.2	1920	US-10-120-198B-1	Sequence 1, Appli
2	1249	51.2	6834	US-10-120-198B-5	Sequence 5, Appli
3	1015.5	41.6	7654	US-10-006-773-1	Sequence 1, Appli
4	1015.5	41.6	7654	US-10-006-771A-1	Sequence 1, Appli
5	994	40.7	771	US-10-013-173-38	Sequence 38, Appli
6	994	40.7	771	US-10-150-762-38	Sequence 38, Appli
7	994	40.7	771	US-10-244-821-38	Sequence 38, Appli
8	981.5	40.2	909	US-09-887-853-1	Sequence 1, Appli
9	969.5	39.7	909	US-09-683-547-1	Sequence 1, Appli
10	930	38.1	1350	US-08-812-393A-1	Sequence 1, Appli
11	928	38.0	1350	US-09-774-681-1	Sequence 1, Appli
12	917.5	37.6	1515	US-10-239-656-78	Sequence 78, Appli
13	913	37.4	816	US-09-766-543-9	Sequence 9, Appli
14	903.5	37.0	723	US-10-127-890-90	Sequence 90, Appli
15	903.5	37.0	723	US-10-717-243-90	Sequence 90, Appli
16	901.5	36.9	804	US-10-071-485-1	Sequence 1, Appli
17	897	36.7	840	US-09-766-543-11	Sequence 11, Appli
18	888.5	36.4	1626	US-10-071-485-84	Sequence 84, Appli
19	888.5	36.4	2133	US-10-071-485-89	Sequence 89, Appli
20	886.5	36.3	1509	US-10-239-656-74	Sequence 74, Appli
21	875.5	35.9	10511	US-10-059-261-109	Sequence 109, Appli
22	875.5	35.9	10511	US-10-627-649-109	Sequence 109, Appli
23	872.5	35.7	729	US-10-879-994-9	Sequence 9, Appli
24	872.5	35.7	729	US-10-610-452-9	Sequence 9, Appli
25	862	35.3	1869	US-10-378-832A-1	Sequence 1, Appli
26	860.5	35.0	761	US-10-354-246-4	Sequence 4, Appli
27	854.5	35.0	1509	US-10-239-656-76	Sequence 76, Appli
28	854	35.0	777	US-10-096-246-7	Sequence 5, Appli
29	854	35.0	777	US-10-096-246-7	Sequence 7, Appli
30	848.5	34.8	1497	US-10-239-656-72	Sequence 72, Appli
31	848	34.7	702	US-10-071-485-92	Sequence 92, Appli
32	846	34.7	777	US-10-096-246-6	Sequence 6, Appli
33	841.5	34.5	756	US-09-984-186-17	Sequence 17, Appli
34	841.5	34.5	756	US-10-237-667-17	Sequence 17, Appli
35	841.5	34.5	756	US-10-237-667-17	Sequence 17, Appli
36	841.5	34.5	756	US-10-237-666-17	Sequence 17, Appli
37	841.5	34.5	756	US-10-237-871-17	Sequence 17, Appli
38	841.5	34.5	756	US-10-237-624-17	Sequence 17, Appli
39	841.5	34.5	756	US-10-702-536-17	Sequence 17, Appli
40	841.5	34.5	756	US-10-702-536-17	Sequence 17, Appli
41	839	34.4	921	US-10-363-349-12	Sequence 12, Appli
42	838	34.3	822	US-09-749-873-108	Sequence 108, Appli
43	833.5	34.1	2090	US-10-104-522-6	Sequence 6, Appli
44	833.5	34.1	2090	US-10-060-585-6	Sequence 6, Appli
45	833.5	34.1	2090	US-10-334-235-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-10-120-198B-1
Sequence 1, Application US/10120198B
Publication No. US20030215427A1
GENERAL INFORMATION:
APPLICANT: Jensen, Michael
TITLE OF INVENTION: CE7-SPECIFIC REDIRECTED IMMUNE CELLS
FILE REFERENCE: 1954-337
CURRENT APPLICATION NUMBER: US/10/120, 198B
PRIOR FILING DATE: 2002-04-11
PRIOR APPLICATION NUMBER: 60/282, 859
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1920
TYPE: DNA
ORGANISM: artificial sequence
FEATURE:
OTHER INFORMATION: mouse-human chimera

```
FEATURE:
; NAME/KEY: CDS
; LOCATION: (14) (1906)
; OTHER INFORMATION: scrvfc construct
US-10-120-198B-1

Alignment Scores:
Pred. No.: 8,64e-121 Length: 1920
Score: 1249.00 Matches: 276
Percent Similarity: 54.60% Conservative: 62
Best Local Similarity: 44.59% Mismatches: 96
Query Match: 51.17% Indels: 186
DB: 17 Gaps: 12

US-09-596-774-6 (1-461) x US-10-120-198B-1 (1-1920)

QY 19 GlnIleGlnLeuValGlnSerGlyProGluLeuLysLysProGlyGluThrValLysIle 38
DB 80 CAGGTCACACCTGCACAGCTGGGGCTGAACTGTAAGCCCTGGGGCTTCAGTGAAGCTG 139
QY 39 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValLysIle 58
DB 140 TCCTCAAGGCTTCCTGGCTACCTTCACCGGCTACGTGACTGGGTGAAGCAGG 199
QY 59 ProGlyGlnGlyLeuLysTyrMetGlyTyrIleAsnThrSerThrGlyGluSerThrPhe 78
DB 200 CCTGACATGGGCTTGAGTGATGGAGAGATTATCTCAGCAAGCTGCTAAGTACTAC 259
QY 79 AlaAspAspPheLysGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 98
DB 260 AATGAGAGTTTCAGAGAGAGCCACCTGACTGTAGCAAACTCTCCACACAGCCCTTC 319
QY 99 LeuGlnIleAsnLeuLysSerGlyLysAspMetAlaThrTyrPheCysAlaArgTyrGlu 118
DB 320 ATGCACCTCAGCGGCTTACATCTGAGACTCTGCAAGCTATTCTTCTGCAAGA----- 373
QY 119 ValTyrHisGly-----TyrValProTyrTyrGlyGlnGlyThrValThrVal 135
DB 374 GATTACTACGGTACTAGCTACCACTTGTACTACGGGCAAGGACCACTCTCACAGTC 433
QY 136 SerSerGlyGlyGlySerGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 155
DB 434 TCCCGAGAGGTGGGTGAGTGGAGGTGGCGGATCCGGTGGCGAGTGAAGCATCCAG 493
QY 156 LeuThrGlnSerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCys 175
DB 494 ATGACACATCTTCATCTCTCTTTCTCTATCTCTAGAGACAGAGTCACATTACTTC 553
QY 176 LysAlaSerGlnAspValTyrAsnAlaValAlaTyrTyrGlnGlnLysProGlyGlnSer 195
DB 554 AAGGCTATGAGACATATAATATCGGTAGCCGTGATCAGACAGACACCGAATAATCT 613
QY 196 ProLysLeuLeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThr 215
DB 614 CTAAGGCTCTTAATATCTGTCGCAACCAATTGTGTACTGGGGTTCCTTCAAGATTAGT 673
QY 216 GlySerGlySerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeu 235
DB 674 GCGAGTGATCTGAAAGGATTACACTCTCACATTACAGCTCTCAGGCTGAAGATTTT 733
QY 236 AlaValTyrPheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyTyrLys 255
DB 734 GCTACTATATTAATCTGCAACATATTGAGTACTCAATTCAGTTCCGCTCGGGAGCAGAG 793
QY 256 LeuGluIleLys----- 259
DB 794 CTGAGATCAAAAGTAGAACCAAAATCTTTCGACAAAACACACATCCCAACGCTGCCA 853
QY 259 ----- 259
DB 854 GCACCTGAAGCTCTGGGGGAGCCGTAGTCTTCTCTTCCCCCAAAACCAAGAGACACC 913
QY 260 -----AlaLeuGluIleSer----- 264
```

```
DB 914 CTATGATCTCCGGAGACCCCTGAGTCAATGCTGTGTGGAGCTGAGCCAGAAAC 973
QY 264 ----- 264
DB 974 CCGAGGTCAAAGTTCAACTGGTACTGTGACGGCGTGGAGGTCATATGCAAGCAAG 1033
QY 265 -----AsnSerValMetTyrPheSerSerValValProValLeuGln 278
DB 1034 CCGGGGAGAGAGAGTCAACACAGCAGTACCGTGTGTGCTCAGCCCTCAGCCCTGAC 1093
QY 279 -----LysValAsnSerThrThrThrLysPyr 287
DB 1094 CAGAGCTGGCTGATGAGCAAGAGAGTACAGTGCAGAGTCTCCAAAGCCCTCCAGCC 1153
QY 287 oValLeuArgThrProSerProValHisProThrGlyThrSerGlnProGlnArgProG 307
DB 1154 CCGATCGAGAAACATCTCCA---AAGTCAAAGGAGACCCGAGAAACACAGGTGTAC 1210
QY 307 uAspCys----- 309
DB 1211 ACCCTGCACCATCAGAGATGAGCTGATCCAGAAACAGGTACGCTGACCTGTGTC 1270
QY 310 -----ArgProArgGlySerValLysGly 317
DB 1271 AAAGGCTTATATCCAGCAGCATCGCGGTGAGTGGAGAGCATGGCGACCGAGAAC 1330
QY 317 YThrGlyLeuAspPheLeuGluAspProLysLeu---CysTyrLeuLeuAspGlyIle 336
DB 1331 AACCT-----ACAAGACCAACGCTCCGCTGTGATCC--GACGGCTCCTT 1374
QY 336 uPheIleTyr----- 339
DB 1375 CTTCCTTACAGCAAGCTACCGTGCACAGAGAGTGGCAGCAGGGAGACGTCCTTC 1434
QY 339 ----- 339
DB 1435 ATGTCGCTGATGATGAGGCTTGCAACACATACAGCAGAAAGCCTCCTGTC 1494
QY 340 -----GlyValIleIle----- 343
DB 1495 TCCCGGGAATAATGCCCTGATTTGTGCGAGGGCGTGGCGGCTCTCTTTCATTGG 1554
QY 344 -ThrAlaLeuTyrLeuArgAlaLysPheSerArgSerAlaGluThrAlaAlaAsnLeuG 363
DB 1555 GCTAGGCAATTTCTTACAGAGTAACTTCAAGAGAGCGAGAGCGCCCGCGTACAGCA 1614
QY 363 nAspProAsnGlnLeuTyrAsnGluLeuAsnLeuGlyArgArgGluGluTyrAspValle 383
DB 1615 GGCCAGAAACAGCTCTATTAAGAGCTCAATCTAGAGACGAAGAGAGATGATGATTTT 1674
QY 383 uGluLysLysArgAlaArgAspProGluMetGlyGlyLysGlnGlnArgArgAsnPr 403
DB 1675 GACCAAGAGACGTGGCGGACCTTGAATGGGGGGAAG---CCGAGAGAGAAAGAACCC 1731
QY 403 oGlnGlnGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaGluAlaTyrSerGluI 423
DB 1732 TCGAGAGGCTCTTCAATGAATGAGTGCAGAAAGTAAAGTGGGAGGCTTACAGTGAAGT 1791
QY 423 eGlyThrLysGlyLysArgArgArgGlyLysGlyHisAspGlyLeuTyrGlnGlyLeu 443
DB 1792 TGGGATGAAGAGGCGAGCGCGGAGGGGCAAGGGGCGAGTGGCTTTCACAGGCTCAG 1851
QY 443 rThrAlaThrLysAspThrTyrAsnAlaLeuHisMetGlnThrLeuAlaProArg 461
DB 1852 TACAGCCACCAAGAGACCTTACAGCGCTTCAATGACGGCTTGGCCCTTGGC 1906

RESULT 2
US-10-120-198B-5
; Sequence 5, Application US/10120198B
; Publication No. US20030215427A1
; GENERAL INFORMATION:
; APPLICANT: Jensen, Michael
```

TITLE OF INVENTION: CE7-SPECIFIC REDIRECTED IMMUNE CELLS
 FILE REFERENCE: 1954-337
 CURRENT APPLICATION NUMBER: US/10/120,198B
 CURRENT FILING DATE: 2002-04-11
 PRIOR APPLICATION NUMBER: 60/282,859
 PRIOR FILING DATE: 2001-04-11
 NUMBER OF SEQ ID NOS: 11
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 5
 LENGTH: 6834
 TYPE: DNA
 ORGANISM: artificial sequence
 FEATURE:
 OTHER INFORMATION: plasmid construct
 US-10-120-198B-5

Alignment Scores:
 Pred. No.: 4,87e-120 Length: 6834
 Score: 1249.00 Matches: 276
 Percent Similarity: 54.60% Conservative: 62
 Best Local Similarity: 44.59% Mismatches: 96
 Query Match: 51.17% Indels: 186
 Gaps: 12

US-09-596-774-6 (1-461) x US-10-120-198B-5 (1-6834)

QY 19 GlnIleGlnLeuValGlnSerGlyProGluLeuValSerProGlyGluThrValIle 38
 DB 1878 CAGGTCACACGACGACGCTGGGCTGAACCTGGAAGCCGGGGCTTCAAGTGAAGCTG 1937
 QY 39 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValIleGlnAla 58
 DB 1938 TCCGCAAGGCTTCTGCTACACCTTCAACGGCTACCTGATGCACTGAGTGAAGAGAG 1997
 QY 59 ProGlyGlnGlyLeuValTyrPheGlyTyrPheLeuThrSerThrGlyGluSerThrPhe 78
 DB 1998 CCTGACACATGCTGCTGATGATGAGAGATTAACTCTAGCAACGGCTGCTACTACTAC 2057
 QY 79 AlaSerPhePheValGlyArgPheSerPheSerLeuGluThrSerAlaAsnThrAlaTyr 98
 DB 2058 AATAGAGGTTCAAGAGCAAGCAACCTGACTGAGCAATCTCTCAGCAAGCCCTTC 2117
 QY 99 LeuGlnIleAsnLeuValSerGlyLysPheAlaThrTyrPheCysAlaArgTyrPhe 118
 DB 2118 ATGCACTCAGCGGCTGATCTGAGAGCTCTGAGCTATTTCTGCAAGA----- 2171
 QY 119 ValTyrHisGly-----TyrValProTyrTyrGlyGlnGlyTyrThrValThrVal 135
 DB 2172 GATTACTACGGTACTAGCTACCACTTGTACTAGTGGGCAAGGACCACTCTCAGCTC 2231
 QY 136 SerSerGlyGlyGlyGlySerGlyGlyGlySerGlyGlyGlySerAspIleGln 155
 DB 2232 TCCCTCAGAGGCTGCGGTAGTGGAGTGGCGGATCCGGTGGCGGAGTGAACATTCAG 2291
 QY 156 LeuThrGlnSerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCys 175
 DB 2292 ATGACACATTTCTATCTCTCTTCTGATCTCTAGAGAGACAGAGTACCATTTACTTGC 2351
 QY 176 LysAlaSerGlnAspValTyrAsnAlaValAlaTyrTyrGlnGlnLysProGlyGlnSer 195
 DB 2352 AAGGCTAATGAACATTAATTAATCGGTACTGATCAGACAGACACCGAATAATCT 2411
 QY 196 ProLysLeuLeuIleTyrSerHisSerSerArgTyrThrGlyValProSerArgPheThr 215
 DB 2412 CCTAGGCTCTTAATATCTGCTGCAACAAATTGGTACTGGGGTCTTCAAGATTCTG 2471
 QY 216 GlySerGlySerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeu 235
 DB 2472 GCGAGTGAATCTGGAAGAGATTACCTCTCAGCATTTACAGTCTTCAAGGCTGAAGATT 2531
 QY 236 AlaValTyrPheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyTyrLys 255
 DB 2532 GCTACTATTACTCTCAACATATTGAGTACTCATTCAGCTTGGGCTGGGAGACAGAG 2591

QY 256 LeuGlnIleLys----- 259
 DB 2592 CTCGAGATCAAGTAAGAACCCAAATCTTCTGACAAAACCTACACATGCCACCGGCCCA 2651
 QY 259 ----- 259
 DB 2652 GCACCTGAATCTCGGGGGGACCGTCACTTCTCTTCCCCCCCCAAAACCAAGACACC 2711
 QY 260 -----AlaLeuGlnIleSer----- 264
 DB 2712 CTGATGATCTCCCGACCCCTGAGGTCAATGCGTGTGTGAGACGTGAGCAGACAGAGAC 2771
 QY 264 ----- 264
 DB 2772 CCTGAGTCAAGTTCACTGCTGATCGGACCGCTGAGGTGCATTAATGCCAAGCAAG 2831
 QY 265 -----AsnSerValMetTyrPheSerSerValProValLeuGln 278
 DB 2832 CCGCGAGAGAGACAGTACACAGACAGTACCGTGTGCTGCTCCTCAGCTCTGAC 2891
 QY 279 -----LysValAsnSerThrThr-ThrLysPr 287
 DB 2892 CAGACTGGCTGAATGGCAAGAGTACAAAGTGCAGAGTCTCCAAAGCCCTCCAGCC 2951
 QY 287 ovalLeuArgThrProSerProValHisProThrGlyTyrSerGlnProGlnArgProGly 307
 DB 2952 CCATTCGAGAAACATTTCCA--AGCCAAAGGCGACCCCGAAGAACACAGGTGTAC 3008
 QY 307 uAspCys----- 309
 DB 3009 ACCCTGCACATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3068
 QY 310 -----ArgProArgGlySerValLysGly 317
 DB 3069 AAAGCTTCTATCCAGAGCATGCGCTGAGTGGGAGACAAATGGCGACCCGAGAAC 3128
 QY 317 YThrGlyLeuAspPheLeuGluAspProLysLeu-----CysTyrLeuLeuAspGlyIle 336
 DB 3129 AACT-----ACAAAGACAGGCTCTCCGCTGAGCTC-GACGGCTCTT 3172
 QY 336 uPheIleTyr----- 339
 DB 3173 CTTCCTTACAGCAAGCTCACCGTGAACAAGACAGGTCGACAGAGGAAAGTCTTCTC 3232
 QY 339 ----- 339
 DB 3233 ATGCTCCGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3292
 QY 340 -----GlyValIleIle----- 343
 DB 3293 TCCCGGGAATGCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3352
 QY 344 -ThrAlaLeuTyrLeuArgAlaLysPheSerArgSerAlaGluThrAlaAlaAsnLeuGly 363
 DB 3353 GCTAGGATCTTCTTCAAGAGTGAAGTTTACAGAGAGGCAAGCCCGCGGTACACAGCA 3412
 QY 363 nAspProAsnGlnLeuTyrAsnGlnLeuAsnLeuGlyArgArgGluTyrAspValLe 383
 DB 3413 GGGCAGAACCGCTCTATTAACGCTCAATCTGAGCAAGAGAGAGTGCATGTTT 3472
 QY 383 uGluLysLysArgAlaArgAspProGluMetGlyLysGlnGlnLysArgArgAsnPr 403
 DB 3473 GGAACAGAGAGTGGCGGAGCCCTGAGATGGGGGAAAG---CCGAGAAAGAAAGACC 3529
 QY 403 oGlnGluGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaGluAlaTyrSerGly 423
 DB 3530 TCAGAGAGGCTCTTACATGAATGAGCAAGAAAGATGAATGGCGGCTTACATGATGAT 3589
 QY 423 eGlyThrLysGlyGluArgArgArgGlyLysGlyHisAspGlyLeuTyrGlnGlyLeu 443
 DB 3590 TGGGATGAAGAGCGAGCGCGGAGGGGCAAGGAGGAGTGCCTTTACAGGCTCTCAG 3649

QY 443 rThralaThrLyseAspThrTyraAspAlaLeuHismetGlnThrLeuAlaProArg 461
 Db 3650 TAAAGCCACCAAGAGACCTTAACAGCCCTTACATAGAGGCCCTCCCTCGC 3704
 RESULT 3
 US-10-006-773-1
 ; Sequence 1, Application US/10006773
 ; Publication No. US20020132983A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Junghans, Richard P.
 ; TITLE OF INVENTION: Antibodies as Chimeric Effector Cell Receptors Against Tumor Anti-
 ; FILE REFERENCE: 003
 ; CURRENT APPLICATION NUMBER: US/10/006, 773
 ; PRIORITY FILING DATE: 2001-12-10
 ; PRIOR APPLICATION NUMBER: 60/250, 089
 ; PRIORITY FILING DATE: 2000-11-30
 ; NUMBER OF SEQ ID NOS: 19
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 1
 ; LENGTH: 7654
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens and Mus sp.
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (2428)..(3759)
 ; OTHER INFORMATION: Chimeric IgTcR sequence contained in retroviral vector. Retrovi-
 ; OTHER INFORMATION: al vector sequence (non-coding regions) are incidental to the inv-
 ; OTHER INFORMATION: ention. The translated (coding region) is relevant to the invent-
 ; US-10-006-773-1
 Alignment Scores:
 Pred. No.: 1,996-95 Length: 7654
 Score: 1015.50 Matches: 225
 Percent Similarity: 57.74% Conservative: 25
 Best Local Similarity: 51.96% Mismatches: 58
 Query Match: 41.60% Indels: 125
 Gaps: 8
 DB: 13
 US-09-596-774-6 (1-461) x US-10-006-773-1 (1-7654)
 QY 149 GlyGlyGlySerAspAlaGlnLeuThrGlnSerHisLysPheLeuSerThrSerValGly 168
 Db 2473 GGTGCCACTCCGACATCCACTGACCTGACCCAGCCCAAGCCCTGACGCCAGGGTGGT 2532
 QY 169 AsparGValSerIleThrCysValAspSerGlnAspValTyrAsnAlaValAlaTyrTyr 188
 Db 2533 GACAGAGTGAACCATCCTGTAAAGCCAGTCAAGATGTGGTACTTCTGTAGCTTGGTAC 2592
 QY 189 GlnGlnLysProGlyGlnSerProLysLeuLeuIleTyrSerAlaSerSerArgTyrThr 208
 Db 2593 CAGCAAGAGCCAGGTAAAGCTCCAAAGCTGTGATCTTAATGACATCCACCCGCGACACT 2652
 QY 209 G1ValAlaProSerArgPheThrGlySerGlySerGlyProAspPheThrPheThrIleSer 228
 Db 2653 GGTTGGCCAGCAATTCAGGGGTAAAGGTCGAGTAAAGGTCGACATTCACCTTCACTCAATGAGC 2712
 QY 229 SerValGlnAlaGlnAspLeuAlaValTyrPheCysGlnGlnHis-----PheArgThr 246
 Db 2713 AGCTTCAGCCAGAGACATCCGACCTACTAGCCAGCAATAATAGCTCTATCGG--- 2769
 QY 247 ProPheThrPheGlySerGlyThrLysLeuGlnIleLys----- 259
 Db 2770 -----TCGTTCCGCCCAAGAGGACCAAGGTGAANTCAAAAGAGGTGGCTCAGATGGGT 2823
 QY 259 ----- 259
 Db 2824 GGATCCGGCTCTGTGGCTCAGATCGAGATCCAACTGGTGAGAGCGGTGAGAGTGT 2883
 QY 260 -----AlaLeuGlnIleSerAsnSerVal----- 267
 Db 2884 GTGCACACTGGCGCGCTGCGCTGTCTCTGCTCGCATCTGGATTCGATTCACACACA 2943

QY 267 ----- 267
 Db 2944 TATTGATGATGTTGGTGAGACAGGACACTGTGAAGAAAGTCTTGATGTGATTTGAGAAATT 3003
 QY 267 ----- 267
 Db 3004 CATCCAGATAGACATGACGATTAACTATGCGCCGCTCTTAAAGATATGATTTACAAATATCG 3063
 QY 267 ----- 267
 Db 3064 CGAGACAAAGCCAGAAACACATTGTTCTCGAAATGACAGCCTGAGACCCGAAACACC 3123
 QY 268 -----MetTyrPheSerSer 272
 Db 3124 GGGGTCTATTGTTGTGCAACCTTACTTCGGCTTCCCTGTTGGTTATTTGGGGCCAA 3183
 QY 273 ValValProValLeuGlnLysValAsnSerThrThrLysProValLeuAlaGlyThrPro 292
 Db 3184 GGGACCCCGGTACCGTCTTCACTGCTTAAGCCCAACAGCGCCAGCGCCGACCA 3243
 QY 293 SerProValHisProThrGlyThrSerGlnPro-----GlnArgProGlnAspCysArg 310
 Db 3244 ACAACGGCG---CCACCATCGCGTGCAGCCCTGTCCCTGCGCCAGAGCGCGCTCG 3300
 QY 311 ProArg-----GlySerValLysGlyThrGlyLysAspPheLeuGlnAspProLysLeu 328
 Db 3301 CAGCGCGGGGGGGCGAGTGCACAGAGAGGGGCTGAGCTTGGCCCTGATCCCAAACTC 3360
 QY 329 CysTyrLeuLeuAspGlyLysLeuPheIleTyrGlyValIleIleThrAlaLeuTyrLeu 348
 Db 3361 TGCTACCTGCTGATGAGATCCCTTCACTATGATGATATCTCACTGCTTGTTCCTG 3420
 QY 349 ArgAlaLysPheSerLysSerAlaGlnTyrAlaAlaLeuLeuGlnAspProAsnGlnLeu 368
 Db 3421 AGAGTAAAGTTACAGAGAGCGGAGCCCGCGCTTACACAGAGCGCCAGAACCACTC 3480
 QY 369 TyrAsnGlnLeuAsnLeuGlnLysArgGlnGluTyrAspValLeuGlnLysArgAla 388
 Db 3481 TATTAAGACTCAATTTAGACAGAGAGAGAGAGATGTTTGACAAAGAGAGCTGGC 3540
 QY 389 ArgAspProGlnMetGlyGlyLysGlnGlnArgArgAspProGlnGlnGlyValTyr 408
 Db 3541 CGGACCTCGAATGGGGGGAAG---CCAGAGAGAAAGAACCTCAGAGAGGCTGTAC 3597
 QY 409 AsnAlaLeuGlnLysAspLysMetAlaGlnAlaTyrSerGlnIleGlyThrLysGlyGlu 428
 Db 3598 AATGAACCTGCAAGAAAGATTAAGTGGCGGAGGCTTACAGTGAATGGATGAAGAGCGAG 3657
 QY 429 ArgArgArgGlyLysGlyHisAspGlyLeuTyrGlnGlyLysSerThrAlaThrLysAsp 448
 Db 3658 CCGCGAGGGGCAAGAGGGGACAGATGGCTTTTACAGGGTCTCAGTACAGCCAAAGGANC 3717
 QY 449 ThrTyrAspAlaLeuHisMetGlnThrLeuAlaProArg 461
 Db 3718 ACTTACGAGCGCTTACATGACAGGCCCTGCCCTCGC 3756
 RESULT 4
 US-10-006-771A-1
 ; Sequence 1, Application US/10006771A
 ; Publication No. US20020165360A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Junghans, Richard P.
 ; TITLE OF INVENTION: Chimeric Effector Cell Receptors Against Carcinoembryonic Antigen
 ; FILE REFERENCE: 002
 ; CURRENT APPLICATION NUMBER: US/10/006, 771A
 ; PRIORITY FILING DATE: 2002-06-04
 ; PRIOR APPLICATION NUMBER: 60/250, 090
 ; PRIORITY FILING DATE: 2000-11-30
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 1
 ; LENGTH: 7654
 ; TYPE: DNA

ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (2428)..(3759)
 OTHER INFORMATION: Chimeric IgTcR sequence contained in retroviral vector. Retroviral
 OTHER INFORMATION: at vector sequence (non-coding regions) are incidental to the inv
 OTHER INFORMATION: ention. The translated (coding region) is relevant to the invent
 OTHER INFORMATION: ion. (pertinent to Figure 3.)
 US-10-006-771A-1

Alignment Scores:

Pred. No.:	1,99e-95	Length:	7654
Score:	1015.50	Matches:	225
Percent Similarity:	57.74%	Conservative:	25
Best local Similarity:	51.96%	Mismatches:	58
Query Match:	41.60%	Indels:	125
DB:	13	Gaps:	8

US-09-596-774-6 (1-461) x US-10-006-771A-1 (1-7654)

QY 149 G1yglYglYserAapIleGlnleuThrgInserHslYsPheleuSerThrSerValGly 168
 Db 2473 GGTGTCCACCTCCGACATCCAGCTGACCCAGAGCCCAAGCAGCTGAGCGCGTGGT 2532
 QY 169 AspArgValSerIleThrCysAlaAlaSerGlnAspValTyrAsnAlaValAlaTyrTyr 188
 Db 2533 GACGAGTGCACATCCTGTAGGCCAGTACGATGTGGTACTTCTGTAGCTTGGTAC 2592
 QY 189 GlnGlnIlyProGlyGlnSerProIlyleuLeuIleTyrSerAlaSerSerArgTyrThr 208
 Db 2593 CAGCAGAGGCGAGGTAAGGCTCCAAAGCTGATCTACTGAGACATCAGCCGCGACACT 2652
 QY 209 G1yAlaIProserArgPheThrGlySerGlySerGlyProAspPheThrPheThrIleSer 228
 Db 2653 GGTGTGCCAAGCAGATTCAGGCGTACCGGTACCGACTTCACTTCACTTCACTTCACT 2712
 QY 229 SerValGlnAlaGlnAspLeuAlaValTyrPheCysGlnGlnIle-----PheArgThr 246
 Db 2713 AGCTTCCAGCCAGAGGACATGCGACCTTACTGCTCCAGCATATAGCTTATCG--- 2769
 QY 247 ProPheThrPheGlySerGlyThrIlyleuGlnIlyleYs----- 259
 Db 2770 -----TCGTTCCGCAAGGAGGACCAAGGTGAATCAAAAGAGGTGCTCAGATTCGGGT 2823
 QY 259 ----- 259
 Db 2824 GATCCGCGCTGTGCTGCTCAGATCCGAGGTCCAACTGTGAGAGCGGTGAGGTGT 2883
 QY 260 -----AlaLeuGlnIleSerAsnSerVal----- 267
 Db 2884 GTGCACCTGCGCGGTCCCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2943
 QY 267 ----- 267
 Db 2944 TATTGATGAGTTGGTGAGAGCAGGACCTGAAAAGGTCTTGAAGTATGAGAAAT 3003
 QY 267 ----- 267
 Db 3004 CATCAGATAGACAGTACGATTAAGTATGCGGCTCTTAAGAGATAGATTTACAATATCG 3063
 QY 267 ----- 267
 Db 3064 CGAGACAAACCGCAAGAACACATTGTTCTGCAATAGACAGCTGAGACCGCAAGACACC 3123
 QY 268 -----MetTyrPheSerSer 272
 Db 3124 GGGGTATATTTTGTGCAAGCTTACTTGGGCTTCCCTGTTTGTATTTGGGGCCAA 3183
 QY 273 ValValIProValIleuGlnIlyValAsnSerThrThrIlyIProValIleuArgThrPro 292
 Db 3184 GGAACCCCGGACCGCTCCAGTGAAGCCCAAGACGAGCGCCGCGACCA 3243
 QY 293 SerProValIHisProThrGlyThrSerGlnPro-----GlnArgProGluAspCysArg 310

Db 3244 ACACCGGCG---CCACCATCGCGTCGAGCCCTGTCCTGCGCCCGCAGAGCGCGCTCG 3300
 QY 311 ProArg-----GlySerValIyglYThrGlyIleuAspPheLeuGlnAspProIlyleu 328
 Db 3301 CCACCGGCGGGGGGCGAGGTCACACAGAGGGGCTGAGCTTGGCTGGATCCCAACTC 3360
 QY 329 CysTyrIleuAspArgIlyIleuPheIleTyrGlyValIleIleThrAlaLeuTyrIleu 348
 Db 3361 TGTACTCTGCTGATGAGATCTCTTCACTATGATGTGTCTTCACTGCTTGTCTG 3420
 QY 349 ArgAlaIlyPheSerArgSerAlaGluThrAlaIAsnIleuGlnAspProAsnGlnleu 368
 Db 3421 AGAGTGAAGTTCAACAGAGAGGAGAGAGCCCGCTGACAGAGGCGCAAGAACAGCTC 3480
 QY 369 TyrAsnGlnIleuAsnIleuGlyArgArgGluGluTyrIlyIleuValIleuGlyIlyAla 388
 Db 3481 TATACAGAGCTCAATCTAGAGCGAAGAGAGATCAATGTTTGGACMAAGAGCTGGC 3540
 QY 389 ArgAspProGluMetGlyIlyGlnGlnIlyArgArgAsnProGlnGluIlyValTyr 408
 Db 3541 CGGACCTTGATGAGTGGGGGAAAG---CCGAAAGAAAGAACCTTCAAGAAAGCTGTAC 3597
 QY 409 AsnAlaIleuGlnIlyAspIlyMetAlaGluAlaTyrSerGlnIlyGlyThrIlyGlyIly 428
 Db 3598 AATGACCTGCAAGAAATAGATGCGGAGGCTTACAGTGAATTTGGATGAAGGCGAG 3657
 QY 429 ArgArgArgGlyIlyGlyIlyIlyAspGlyIlyIlyGlnIlyIleuSerThrAlaThrIlyAsp 448
 Db 3658 CGCGGAGGGGCAAGGGGCGAGATGCGCTTACAGAGGTCTCACTAGTACAGCCAAAGGAC 3717
 QY 449 ThrTyrAspAlaIleuHlIleMetGlnThrIleuAlaProArg 461
 Db 3718 ACTTACAGAGCCCTTCACTACAGAGCCCTGCGCCCTGCG 3756

RESULT 5

US-10-013-173-38
 ; Sequence 38, Application US/10013173
 ; Publication No. US20030095977A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Goshorn, Stephen C.
 ; APPLICANT: Graves, Scott Stoll
 ; APPLICANT: Schultz, Joanne Elaine
 ; APPLICANT: Lin, Yukang
 ; APPLICANT: Sanderson, James A.
 ; APPLICANT: Reno, John M.
 ; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
 ; FILE OF INVENTION: METHODS OF USE THEREOF
 ; FILE REFERENCE: 690022.547C1
 ; CURRENT APPLICATION NUMBER: US/10/013.173
 ; CURRENT FILING DATE: 2001-12-07
 ; NUMBER OF SEQ ID NOS: 69
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 38
 ; LENGTH: 771
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: PRIA3 single chain antibody-genomic streptavidin
 ; OTHER INFORMATION: fusion construct
 US-10-013-173-38

Alignment Scores:

Pred. No.:	1,59e-94	Length:	771
Score:	994.00	Matches:	193
Percent Similarity:	82.95%	Conservative:	26
Best local Similarity:	74.81%	Mismatches:	21
Query Match:	40.72%	Indels:	18
DB:	14	Gaps:	5

US-09-596-774-6 (1-461) x US-10-013-173-38 (1-771)
 QY 19 GlnIleGlnIleuValIleuSerGlyProGluIleuIlyIlyProGlyGluIlyIlyValIlyIle 38

US-09-596-774-6 (1-461) x US-09-887-853-1 (1-909)

```

Qy      19  GlnIleGlnLeuValGlnSerGlyProGluLeuLysPheGlyGluThrValIle 38
Db      9  GAGATCCAAATGGTGGACCTGACCTGAGCAAGAGCTGGAGAGACGATCAAGATC 68
Qy      39  SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTrpValIleGlnAla 58
Db      69  TCCTGCAAGGCTTCTGGGTAACTTCAACAATGATGAAATGAACTGGGTGAAGGAGGCT 128
Qy      59  ProGlyGlnGlyLeuLysTrpMetGlyTyrIleAsnThrSerThrGlyGlnSerThrPhe 78
Db      129  CCAGGAAAGGGTTTAAAGTGGATGGCTGATAAACCACTGAGAGGCCAACATAT 188
Qy      79  AlaAspAspPheLysGlyArgPheAspPheSerLeuGlnThrSerAlaAsnThrAlaTyr 98
Db      189  GCTGAAGAAGTTCAAGGACGGTTTGCTTCTTTGGAAACCTTCGACGACCTGCTAT 248
Qy      99  LeuGlnIleAsnAsnLeuLysSerGluAspMetAlaThrTyrPheCysAlaIleTrpGlu 118
Db      249  TTGCAGATCAACAACCTCAAAATAGAGACAGGCTACATATTTCTGTGAAGGCAATT 308
Qy      119  ValTyrHisGlyTyrValProTyrTrpGlyGlnGlyThrValThrValSerSerGly 138
Db      309  ATTACCAAGCGGGGTTTGCTAACTGGGGCCAAAGGACTCTGTCACCTCTCTGCA--- 365
Qy      139  GlyGlyGlySerGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThrGln 158
Db      366  TCGAGCTCCGCGGATCTTCATCTGACGCTTCAGCTCGAGCGATATGCTATGACCCAG 425
Qy      159  SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSer 178
Db      426  TCTCCTAAATTCATGCCACGTGAGGGAGACAGGGTCAGCATCTCCGCAAGGCCAGT 485
Qy      179  GlnAspValTyrAsnAlaValAlaTyrGlnGlnLysProGlyGlnSerProLysLeu 198
Db      486  CAGATGATGATACGCTGCTACCTGCTGATCAACAACCAAGGCAATCTCTCAAACTA 545
Qy      199  LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 218
Db      546  CTGATTTACTGACATCCACCGGACACCTGGAAGTCCCTGATCGCTTCAAGGAGTGA 605
Qy      219  SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 238
Db      606  TCTGGGACAGATTAATACCTCACTCACTCAGAGTGTGAGGCTGAAAGCTGGCACTTCAT 665
Qy      239  PheCysGlnGlnHisPheArgTyrProPheThrPheGlySerGlyTyrIleLysLeuGluIle 258
Db      666  TACTGTACGACAACATTATAGATGCCGTACACGTTCCGAGGGGGACCAAGCTGGAGATA 725
Qy      259  Lys 259
Db      726  AAA 728

```

RESULT 9
US-10-683-547-1

Sequence 1, Application US/10683547
Publication No. US20050058638A1
GENERAL INFORMATION:
APPLICANT: Huston, J.
APPLICANT: Houston, L.L.
APPLICANT: Ring, D.
APPLICANT: Oppermann, H.
TITLE OF INVENTION: BIOSYNTHETIC BINDING PROTEINS FOR IMMUNO-TARGETING
FILE REFERENCE: CIBT-P01-110
CURRENT APPLICATION NUMBER: US/10/683,547
CURRENT FILING DATE: 2003-10-10
PRIOR APPLICATION NUMBER: US/09/558,741
PRIOR FILING DATE: 2000-04-26
PRIOR APPLICATION NUMBER: 07/831,967
PRIOR FILING DATE: 1992-02-06
NUMBER OF SEQ ID NOS: 16
SOFTWARE: PatentIn version 3.1

SEQ ID NO 1
LENGTH: 909
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: 741F8 sfv'
FEATURE:
NAME/KEY: CDS
LOCATION: (3)..(752)
OTHER INFORMATION:
US-10-683-547-1

Alignment Scores:

Pred. No.:	7,47e-92	Length:	909
Score:	969.50	Matches:	183
Percent Similarity:	85.06%	Conservative:	22
Best Local Similarity:	75.93%	Mismatches:	35
Query Match:	39.72%	Indels:	1
DB:	19	Gaps:	1

US-09-596-774-6 (1-461) x US-10-683-547-1 (1-909)

```

Qy      19  GlnIleGlnLeuValGlnSerGlyProGluLeuLysPheGlyGluThrValIle 38
Db      9  GAGATCCAAATGGTGGACCTGACCTGAGCAAGAGCTGGAGAGACGATCAAGATC 68
Qy      39  SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTrpValIleGlnAla 58
Db      69  TCCTGCAAGGCTTCTGGGTAACTTCAACAATGATGAAATGAACTGGGTGAAGGAGGCT 128
Qy      59  ProGlyGlnGlyLeuLysTrpMetGlyTyrIleAsnThrSerThrGlyGlnSerThrPhe 78
Db      129  CCAGGAAAGGGTTTAAAGTGGATGGCTGATAAACCAACCACTGAGAGGCCAACATAT 188
Qy      79  AlaAspAspPheLysGlyArgPheAspPheSerLeuGlnThrSerAlaAsnThrAlaTyr 98
Db      189  GCTGAAGAAGTTCAAGGACGGTTTGCTTCTTTGGAAACCTTCGACGACCTGCTAT 248
Qy      99  LeuGlnIleAsnAsnLeuLysSerGluAspMetAlaThrTyrPheCysAlaIleTrpGlu 118
Db      249  TTGCAGATCAACAACCTCAAAATAGAGACAGGCTACATATTTCTGTGAAGGCAATT 308
Qy      119  ValTyrHisGlyTyrValProTyrTrpGlyGlnGlyThrValThrValSerSerGly 138
Db      309  ATTACCAAGCGGGGTTTGCTAACTGGGGCCAAAGGACTCTGTCACCTCTCTGCA--- 365
Qy      139  GlyGlyGlySerGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThrGln 156
Db      366  TCGAGCTCCGCGGATCTTCATCTGACGCTTCAGCTCGAGCGATATCGTCATGCCAG 425
Qy      159  SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSer 178
Db      426  TCTCCTAAATTCATGCCACGTGAGGGAGACAGGGTCAGCATCTCTCTCAAGGCCAGT 485
Qy      179  GlnAspValTyrAsnAlaValAlaTyrGlnGlnLysProGlyGlnSerProLysLeu 198
Db      486  CAGATGATGATACGCTGCTACCTGCTGATCAACAACCAAGGCAATCTCTCAAACTA 545
Qy      199  LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 218
Db      546  CTGATTTACTGACATCCACCGGACACCTGGAAGTCCCTGATCGCTTCAAGGAGTGA 605
Qy      219  SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 238
Db      606  TCTGGGACAGATTAATACCTCACTCACTCAGAGTGTGAGGCTGAAAGCTGGCACTTCAT 665
Qy      239  PheCysGlnGlnHisPheArgTyrProPheThrPheGlySerGlyTyrIleLysLeuGluIle 258
Db      666  TACTGTACGACAACATTATAGATGCCGTACACGTTCCGAGGGGGACCAAGCTGGAGATA 725
Qy      259  Lys 259
Db      726  AAA 728

```


RESULT 10
US-08-812-393A-1
Sequence 1, Application US/08812393A
Publication No. US20010007152A1
GENERAL INFORMATION:
APPLICANT: SHERMAN, Linda A.
APPLICANT: LUSTGARTEN, Joseph
TITLE OF INVENTION: RECOMBINANT CONSTRUCTS ENCODING
TITLE OF INVENTION: T CELL RECEPTORS SPECIFIC FOR HUMAN HLA-RESTRICTED TUMOR
TITLE OF INVENTION: ANTIGENS
NUMBER OF SEQUENCES: 64
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Avenue, NW, suite 5500
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1888
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/812,393A
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Mureshigse, Kate H
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 31333-20001.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-887-1500
TELEFAX: 202-822-0168
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1350 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
FEATURE:
NAME/KEY: Coding Sequence
LOCATION: 1...1332
OTHER INFORMATION:
US-08-812-393A-1
Alignment Scores:
Pred. No.: 1,82e-87 Length: 1350
Score: 930.00 Matches: 217
Percent Similarity: 59.33% Conservative: 50
Best Local Similarity: 48.22% Mismatches: 145
Query Match: 38.10% Indels: 38
DB: 8 Gaps: 14
US-09-596-774-6 (1-461) x US-08-812-393A-1 (1-1350)
QY 21 GlnleuValGlnSerGlyProGlu-----LeuIlybysProGlyGluThrValIysIle 38
DB 70 CAGCAAGGACGACAGAGTCCCGCATCTTGTTCTGCAAGAGGAGGAGCAAGCAGAGCTC 129
QY 39 SerCysIlybysAlaSerGlyTyrProPheThrIleValIysIleValIysGlnIle 58
DB 130 CAGCTACTTTTC-----ATCTTACAAACAG---GTCCAGTGGTTTACCAAGCT 180
QY 59 ProGlyGlnGlyLeuIlybysTyrPheGlyTyrIleAsnThrSerThrGlyGlnSerThrPhe 78
DB 181 CCGGGGAGAGACCTGCTCAGCTGTTGTACAAATCCTCTGGGACAAAGCAGAGT----- 234

QY 79 AlaAspApePheIySGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 98
DB 235 -----GGGAGACTGACATCCACACAGAGTCAATTAAGAGCTGCGAGCTCT 279
QY 99 LeuGlnIleAsnAsnLeuIlybysSerGluAspMetAlaThrTyrPheCysAlaArgTyrGlu 118
DB 280 TTGCACATTTCTCTCCAGATATCAAGACCTGAGGACCTTATCTGCTGCTCAAAATCT 339
QY 119 ValTyrHisGlyTyrValProTyrTrpGlyGlnGlyThrThrValThrValSerSerGly 138
DB 340 GGAGAGCAATTCGAAAGCTTAACCTTCCGGAAGGCACTTAACCTCTGTTAAATCAGCT 399
QY 139 GlyIlybysSerGlyGlyIlybysSerGlyIlybysSerAspIleGlnLeuThrGln 158
DB 400 GCGGAGGCTGCGGGGGGTGGATCCGGGGGTGGAGGCTCAGAGGCTGAGTCAAGCCCA 459
QY 159 SerHisIlybysPheLeuSerThrSerValGlyAspArgValSerIleThrCysIlybysAlaSer 178
DB 460 AGCCCAAGAAACAAAGTGGAGTAAACAGAGAAAGTGAATTCGATTCGATTCAGACT 519
QY 179 GlnAspValTyrAsnAlaValAlaTyrPheGlnGlnIlybysProGlyGlnSerProIysLeu 198
DB 520 AATAAC---CAACAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 576
QY 199 LeuIlybysSer-----AlaSerSerArgTyrThrGlyValProSerArgPheThr 215
DB 577 ATCATTTATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 636
QY 216 GlySer---GlySerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAsp 234
DB 637 GCTTCAGACCAACCAAGAGAACTTCCCTCATTTCTGAGATTGGTACCCCTCTCAG 696
QY 235 LeuAlaValTyrPheCys-----GlnGlnHisPheArgThrProPheThrPheGly 251
DB 697 ACATGAGTACTTCTGTCGACAGGCTGACAGGACCAAGCAAGAAATTTATTTTCGAT 756
QY 252 SerGlyThrIlybysLeuGlnIleIlybysAlaLeuGlnIlybysSerAsnSerValMetTyrPheSer 271
DB 757 CATGGAACCAAGCTGTCTGCTGACT-----AGTAATCTCATGATGATGATGATGATGAT 807
QY 272 SerValValProValLeuGlnIlybysValAsnSerThrThrIlybysProValLeuArgThr 291
DB 808 CACTTCGTCGCGCTCTCTCCAGAGCAAGCCCAACAGAGC---CCAGCCCGCGGACCA 864
QY 292 ProSerProValHisProThrGlyThrSerGlnProGlnArgProGluAspCysArgPro 311
DB 865 CCAACACCGCGC---CCACCATCGCGTGGCAGCCCTG-----TCCCTGCGGCCA 912
QY 312 ArgGlySerValIlybysGlyThrGlyLeuAspPheLeuGlnAspProIysLeuCysTyrLeu 331
DB 913 TCTAGTTCT-----AGAGATCCCAAACTGTGTAAGCTG 945
QY 332 LeuAspGlyIleLeuPheIleTyrGlyValIleIleThrAlaLeuTyrLeuArgAlaIys 351
DB 946 CTGATGGAATCTCTTCATCTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1005
QY 352 PheSerArgSerAlaGluThrAlaAlaAsnLeuGlnAspProAsnGlnIleTyrAsnGlu 371
DB 1006 TTCAGCAGAGCAGAGAGCCCGCCGCTACAGAGGCGCAGAACCACTCTATTAACAG 1065
QY 372 LeuAsnLeuGlyArgArgGluGluTyrAspValLeuGlnIlybysArgAlaArgAspPro 391
DB 1066 CTCAATCTAGACAGAGAGAGAGAGTACATTTTGGACAGAGAGAGAGAGAGAGAGAGAGAG 1125
QY 392 GlnMetGlyIlybysGlnGlnArgArgArgAsnProGlnGlnIlybysValTyrAsnAlaLeu 411
DB 1126 GAGATGGGGGAGAAAG---CCAGAGAGAGAAAGCCCTCAGAAAGCCCTGTAACAATGAAG 1182
QY 412 GlnIlybysPheMetAlaGluAlaTyrSerGluIleGlyThrIlybysGlyIlybysArgArg 431
DB 1183 CAGAAAGTAAAGATGGCGAGAGCTTACAGTGAATGGAGTGAAGAGAGAGAGAGAGAGAGAG 1242
QY 432 GlyIlybysGlnHisAspGlyLeuTyrGlnGlyLeuSerThrAlaThrIlybysArgThrTyrAsp 451

Db 1243 GGCAGAGGGCAGCATGGCTTTTACAGAGGCTCTCAGTACACCAAGACCTTACGAC 1302
Qy 452 AlaleuHismetGlnThrLeuAlaProArg 461
Db 1303 GCCCTTCATGACGAGCCCTGCCCCCTCGC 1332
RESULT 11
US-09-774-681-1
; Sequence 1, Application US/09774681
; Publication No. US20030208780A1
; GENERAL INFORMATION:
; APPLICANT: Sunol Molecular Corporation
; APPLICANT: Sherman, Linda
; APPLICANT: Lustgarten, Joseph
; TITLE OF INVENTION: RECOMBINANT CONSTRUCTS ENCODING T CELL
; FILE REFERENCE: 31333-20001.01
; CURRENT APPLICATION NUMBER: US/09/774,681
; PRIOR FILING DATE: 2001-02-01
; PRIOR APPLICATION NUMBER: US 08/812,393
; PRIOR FILING DATE: 1997-03-05
; PRIOR APPLICATION NUMBER: US 60/012,845
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1350
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Polynucleotide derivative of effective T cell
; NAME/KEY: CDS
; LOCATION: (1)...(1350)
US-09-774-681-1
Alignment Scores:
Pred. No.: 2,956-87 Length: 1350
Score: 928.00 Matches: 217
Percent Similarity: 59.33% Conservative: 50
Best Local Similarity: 48.22% Mismatches: 145
Query Match: 38.02% Indels: 38
Gaps: 14
US-09-596-774-6 (1-461) x US-09-774-681-1 (1-1350)
Qy 21 GlnLeuValGlnSerGlyProGlu-----LeuYsLysProGlyGluThrValLysIle 38
Db 70 CAGCAAGTGCAGCAGATCCCGCATCTTGGTTCGACGAGGGGGAACGACAGACTC 129
Qy 39 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTrpValLysGlnAla 58
Db 130 CAGGTGAGCTTTTCC-----ATCTTACAAACAG--GTGAGGTGTTTACCAACGT 180
Qy 59 ProGlyGlnGlyLeuYsTrpMetGlyTyrPleAsnThrSerThrGlyGluSerThrPhe 78
Db 181 CTTGGGGGAGACTCTGTCAGCCTGTTTCAATCTTCTGGGACAAAGCAAGCAAGT----- 234
Qy 79 AlaAspAspPheLysGlyArgPheAspPheSerLeuGlnThrSerAlaAsnThrAlaTyr 98
Db 235 -----GGAGACTGCATGCACAAAGTCACTTAAAGACTCGCAGCTC 279
Qy 99 LeuGlnIleAsnLeuYsSerGluAspMetAlaThrTyrPheCysAlaArgTrpGlu 118
Db 280 TTGCACATTTCTCTCCACATCAGACTCAGGACTTATCTGTGCTCAATCTC 339
Qy 119 ValTyrHisGlyTyrValProTyrTrpGlyGlnGlyThrThrValThrValSerSerGly 138
Db 340 CGAGGAAGCAATGCAAGCAAGTCACTTGGGAAAGGCACTAACTCTCTGTTAAATCAGGT 399
Qy 139 GlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThrGln 158

Db 400 GCGGAGGCTGTGCGGGGGTGATCCGGGGTGAGGCTCAGAGGCTGCAGTACCCAA 459
Qy 159 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSer 178
Db 460 AGCCCAAGAAACAAAGTGGCGATGACACAGAGAAAGGACATGAGCTTAAATCAGACT 519
Qy 179 GlnAspValTyrAsnAlaValAlaTrpTyrGlnGlnLysProGlyGlnSerProLysLeu 198
Db 520 AATAAC---CACACAACATGATCTGTATCGGAGAGACGGGGGATGGGCTGAGGCTG 576
Qy 199 LeuIleTyrSer-----AlaSerSerArgTyrThrGlyValProSerArgPheThr 215
Db 577 ATCCATTAATTCATATGCTGTGCTGCGACGACGAGAAAGAGATTCCTGTATGATCAAG 636
Qy 216 GlySer---GlySerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAsp 234
Db 637 GCGTCCAGACCAAGCAAGGAAAGTCTCCCTCATTTGTGAGTGGGTACCCCTTCAG 696
Qy 235 LeuAlaValTyrPheCys-----GlnGlnHisPheArgThrProPheThrPheGly 251
Db 697 ACATCAGTGTACTTCTGTGCGACGGGTGAGACAGGACCAACGAAGATTATTTTCGGT 756
Qy 252 SerGlyThrLysLeuGlnIleLysAlaLeuGlnIleSerAsnSerValMetTyrPheSer 271
Db 757 CATGAAACCAAGCTGTCTGTCTGACT-----AGTAACTCCATCATGTACTCAGC 807
Qy 272 SerValValProValLeuGlnLysValAsnSerThrThrLysProValLeuArgThr 291
Db 808 CACTTCGTGCGGTCTCTCTGCGACGAGAACCCACACAGACG---CGAGCGCGGAGCA 864
Qy 292 ProSerProValHisProThrGlyThrSerGlnProGlnArgProGluAspCysArgPro 311
Db 865 CCAACACCGCG---CCACACATCGCGTGCAGCCCTCG-----TCCCTGGGCCCA 912
Qy 312 ArgGlySerValLysGlyThrGlyLeuAspPheLeuGluAspProLysLeuCysTyrLeu 331
Db 913 TCTAGTTCT-----AGAGATCCCAACTCTGCTACTCG 945
Qy 332 LeuAspGlyIleLeuPheIleTyrGlyValIleIleThrAlaLeuTyrLeuArgAlaLys 351
Db 946 CTGAGAGGATCTCTTCATCTATGATGTCATTCATGCTGCTTCTGCTGAGATGAAG 1005
Qy 352 PheSerArgSerAlaGlnThrAlaAlaAsnLeuGlnAspProAsnGlnLeuThrAsnGlu 371
Db 1006 TTCAGAGAGGCGCAGACCCCGCGTACCGACGAGGCGCCAGAACCACTCATTAACGAG 1065
Qy 372 LeuAsnLeuGlyValArgGlnGlnIleTyrAspValLeuGlnLysValArgAlaArgAspPro 391
Db 1066 CTCATCTTAGACGAAAGAGAGATGATGTTTGGACAAAGACGTGGCGGGACCT 1125
Qy 392 GlnMetGlyGlyLysGlnGlnIleArgArgArgAsnProGlnGlnGlyValTyrAsnAlaLeu 411
Db 1126 GAGATGGGGGGAAG---CCGAGAAAGAAAGACCTTCGAGAGGCTGTAAAGAACTG 1182
Qy 412 GlnLysAspLysMetAlaGlnAlaTyrSerGlnIleGlyThrLysGlyGlnLysArgArg 431
Db 1183 CAGAAAGATTAAGATGGCGAGGCTTACAGATGAGTTGGAGTGAAGAGCGCCGAGAG 1242
Qy 432 GlyLysGlyHisAspGlyLeuTyrGlnGlyLeuSerThrAlaThrLysAspThrTyrAsp 451
Db 1243 GCGAAGGGGACAGAGGCTTTTACAGGCTCTCAGTACAGCCAAAGACACCTTACGAC 1302
Qy 452 AlaleuHismetGlnThrLeuAlaProArg 461
Db 1303 GCCCTTCATGACGAGCCCTGCCCCCTCGC 1332
RESULT 12
US-10-239-656-78
; Sequence 78, Application US/10239656
; Publication No. US2004003839A1
; GENERAL INFORMATION:
; APPLICANT: KUPFER, PETER
; APPLICANT: RIETHMULLER, GERT

```

1  APPLICANT: LUTTERBUSE, RALF
2  APPLICANT: BORSCHERT, KATRIN
3  APPLICANT: KISCHEL, ROMAN
4  APPLICANT: MAYER, MONIKA
5  APPLICANT: HOMMEISTER, ROBERT
6  TITLE OF INVENTION: MULTIFUNCTIONAL POLYPEPTIDES COMPISING A BINDING SITE*
7  TITLE OF INVENTION: TO AN EPIOTOPE OF THE NG2D RECEPTOR COMPLEX
8  FILE REFERENCE: 029976/0106
9  CURRENT APPLICATION NUMBER: US/10/0239,656
10 CURRENT FILING DATE: 2003-03-06
11 PRIOR APPLICATION NUMBER: PCT/EP01/03414
12 PRIOR FILING DATE: 2001-03-26
13 PRIOR APPLICATION NUMBER: EP 00106467.4
14 PRIOR FILING DATE: 2000-03-24
15 NUMBER OF SEQ ID NOS: 92
16 SOFTWARE: PatentIn Ver. 2.1
17 SEQ ID NO 78
18 LENGTH: 1515
19 TYPE: DNA
20 ORGANISM: Artificial Sequence
21 FEATURE:
22 OTHER INFORMATION: Description of Artificial Sequence: Synthetic 3810xps-
23 US-10-239-656-78 23 Dispecific Single chain fv

```

[illegible]

Db	862	CCAGGAAAGGGCTAGAAATGGAATTGGAGAAATTATTCACAGTATGACAGTAAACAT	921
Oy	117	-----	117
Db	922	ACGCCATCTCTAAAGATAGATTCATCATCTCCAGACACACGCCAAAAATACGCTGAC	981
Oy	118	-----GluVal ¹ TrpHis-----	121
Db	982	CTGCAAATGACCAAGTGAAGTCTGAGACACACGCCCTTTATTACTGTGCAGATTGGGG	1041
Oy	122	-----Gly ¹ TrpVal ¹ Pro ¹ TyrTrpGlyGlnGly ¹ Thr ¹ ThrVal ¹ ThrVal ¹ SerSerGly ¹ Gly ¹	139
Db	1042	CAATGGGGGATCTTGAAGTACTGAGGCGCAAGGACACACGGTCAACCTCTCCTCAGTGGT	1101
Oy	140	GlyGlySerGly ¹ Gly ¹ Gly ¹ SerSerGly ¹ Gly ¹ Gly ¹ SerAsp ¹ IleGlnLeu ¹ ThrGlnSer	159
Db	1102	GGTGATCTCGCGCGCGCGGCTCGCGTGGTGGTGTCTGAGCTCCGATGACACACTCT	1161
Oy	160	HisLysPheLeuSer ¹ ThrSerVal ¹ GlyAspArgValSer ¹ Ile ¹ ThrCysLysValSerGln	179
Db	1162	CCATCTCTCCCTGACCTGACCTGACACGACGAGAGAGGGTCACTAGAGCTGCAAGTCCAGTCAAG	1221
Oy	180	AspVal ¹ TyrAsnAla-----Val ¹ AlaTrp ¹ TyrGlnGlnLysProGly ¹	193
Db	1222	AGTCTGTTAAACAGTGAATAATCAAAAGAACTACTTGGACCTGGTACCGAGAAACCAAGGG	1281
Oy	194	GlnSerProLysLeuLeu ¹ Ile ¹ TyrSer ¹ AlaSerSerArg ¹ Tyr ¹ ThrGlyVal ¹ ProSerArg	213
Db	1282	CAGCTCTTAAACGTGGATCTACTGCGACATCCACTGGGAGAAATGTGGAGTCCCTGATGCG	1341
Oy	214	Phe ¹ ThrGlySerGlySerGly ¹ ProCAspPhe ¹ ThrPhe ¹ Thr ¹ IleSerSerVal ¹ GlnAlaGlu	233
Db	1342	TTTCACAGGACGTGATCTGGAACAGATTTCACCTCAACATCCAGCACTGTGACAGCTGAA	1401
Oy	234	AspLeuAlaVal ¹ TyrPheCysGlnGlnHisPheArg ¹ ThrPro ¹ Phe ¹ ThrPheGlySerGly ¹	253
Db	1402	GACCTGGCAGTTTATTACTGCAAGATGAATTAATGATTATCTCTCAACGTTTCGGGTGGG	1461
Oy	254	ThrLysLeuGlnIleLys	259
Db	1462	ACCAAGCTTGAGATCAAA	1479

```

RESULT 13
US-09-766-543-9
; Sequence 9, Application US/09766543
; Patent No. US20020041865A1
; GENERAL INFORMATION:
; APPLICANT: Austin, Richard
; APPLICANT: Austin, Cheuk S.
; APPLICANT: Ring, David B.
; TITLE OF INVENTION: METHODS FOR TREATING TUMORS
; FILE REFERENCE: P01679.002
; CURRENT APPLICATION NUMBER: US/09/766,543
; CURRENT FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: 60/177,258
; PRIOR FILING DATE: 2000-01-20
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 9
; LENGTH: 816
; TYPE: DNA
; ORGANISM: Artificial Sequence
FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: 520C9
; OTHER INFORMATION: humanized single-chain antibody used in the
; NAME/KEY: CDS
; LOCATION: (7)..(807)
US-09-766-543-9

Alignment Scores:
Pred. No.:      5.6e-86      Length:      816
Score:          913.00      Matches:     167

```

Percent Similarity: 83.61%
 Best Local Similarity: 68.44%
 Query Match: 37.40%
 DB: 9
 Gaps: 2

US-09-596-774-6 (1-461) x US-09-766-543-9 (1-816)

Conservative: 37
 Mismatches: 36
 Indels: 4

QY 19 GlnIleGlnLeuValGlnSerGlyProGluLeuValProGlyValThrValIle 38
 Db 73 GAGATCAACGTGGTGAAGTCTGGGCTGAGTGAAGAACCTGGGCGCTCACTGAAGGTC 132
 QY 39 SerCysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValIle 58
 Db 133 TCCTGCAAGCTTCTGGTACACCTTACCACTAGAACTGAGTGGGTGGACAGGCT 192
 QY 59 ProGlyGlnGlyLeuValTyrMetGlyTyrPheThrSerThrGlyGluSerThrPhe 78
 Db 193 CCTGACCAAGGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 252
 QY 79 AlaAspAspPheLeuValArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 98
 Db 253 GCTGATGACTTCAAGAAAGAGTCAACCACTGACAGACACCTCAGACAGACCTTAC 312
 QY 99 LeuGlnIleAsnAsnLeuValSerGlyLeuAspMetAlaThrTyrPheCysAlaArgTyrPhe 118
 Db 313 ATGGAACCTGAGAGCTGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 366
 QY 119 ValTyrHisGlyTyrValProTyrThrPheGlnGlyThrThrValThrValSerGly 138
 Db 367 ---CAATTTGGTGGTCT---TACTGGGCGCAGGAACTGGTCACTGCTCTCTCAGGT 420
 QY 139 GlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlyGlySerAspIleGlnLeuThrGln 158
 Db 421 GCGCGTGGCTGGGCGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 480
 QY 159 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLeuAlaSer 178
 Db 481 TCTCATCT 540
 QY 179 GlnAspValTyrAsnAlaValAlaTyrPheGlnGlyGlnSerProGlyGlnSerProGlyLeu 198
 Db 541 CAGACATGTTGTAATAGCTTAACCTGATCAGACAGAAACAGGAAACCCCTTAAGCTC 600
 QY 199 LeuIleTyrSerAlaSerSerThrArgTyrThrGlyValProSerArgPheThrGlySerGly 218
 Db 601 CTGATCTACACCATCCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTCAGTTC 660
 QY 219 SerGlyProAspPheThrPheThrIleSerSerValGlnIleGluAspLeuAlaValTyr 238
 Db 661 TCTGGACAGATTTCATCTTCATCAGCATCAGCATCAGCATCAGCATCAGCATCAGCAT 720
 QY 239 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyThrIleLeuGluIle 258
 Db 721 TACTGTCTACAAATATGCTATTTTCCTGACAGCTTCGACCAAGGACAGACAGCTGAGATT 780
 QY 259 LysAlaLeuGlu 262
 Db 781 AAAGATCCGAA 792

RESULT 14
 US-10-127-890-90
 Sequence 90. Application US/10127890
 Publication No. US20030166196A1
 GENERAL INFORMATION:
 APPLICANT: Better, Marc D.
 Studnika, Gary M.
 Carroll, Stephen F.

TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
 Proteins
 NUMBER OF SEQUENCES: 173
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: McAndrews, Held & Malloy, Ltd.
 STREET: 500 West Madison Street, 34th floor

CITY: Chicago
 STATE: Illinois
 COUNTRY: USA
 ZIP: 60661

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/10/127,890
 FILING DATE: 23-Apr-2002
 CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US/08/646,360
 FILING DATE: 13-MAY-1996
 APPLICATION NUMBER: PCT/US94/05348
 FILING DATE: 12-MAY-1994
 APPLICATION NUMBER: US 08/064,691
 FILING DATE: 12-MAY-1993
 APPLICATION NUMBER: US 07/988,430
 FILING DATE: 09-DEC-1992
 APPLICATION NUMBER: US 07/901,707
 FILING DATE: 19-JUN-1992
 APPLICATION NUMBER: US 07/787,567
 FILING DATE: 04-NOV-1991

ATTORNEY/AGENT INFORMATION:
 NAME: McNicholas, Janet M.
 REGISTRATION NUMBER: 32,918
 REFERENCE/DOCKET NUMBER: 200-70.P4

TELECOMMUNICATION INFORMATION:
 TELEPHONE: 312/707-8889
 TELEFAX: 312/707-9155
 TELEX: 650 368-1248

INFORMATION FOR SEQ ID NO: 90:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 723 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA
 SEQUENCE DESCRIPTION: SEQ ID NO: 90:
 US-10-127-890-90

Alignment Scores:
 Pred. No.: 4,73e-85
 Score: 903.50
 Percent Similarity: 82.57%
 Best Local Similarity: 69.29%
 Query Match: 37.01%
 DB: 16
 Gaps: 1

US-09-596-774-6 (1-461) x US-10-127-890-90 (1-723)

QY 19 GlnIleGlnLeuValGlnSerGlyProGluLeuValProGlyValThrValIle 38
 Db 1 GAGATCAACGTGGTGAAGTCTGGGCTGAGTGAAGAACCTGGGCGCTCACTGAAGGTC 60
 QY 39 SerCysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValIle 58
 Db 61 TCCTGCAAGCTTCTGGTACACCTTACCACTAGAACTGAGTGGGTGGACAGGCT 120
 QY 59 ProGlyGlnGlyLeuValTyrMetGlyTyrPheThrSerThrGlyGluSerThrPhe 78
 Db 121 CCTGACCAAGGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 180
 QY 79 AlaAspAspPheLeuValArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 98
 Db 181 GCTGATGACTTCAAGAAAGAGTCAACCACTGACAGACACCTCAGACAGACCTTAC 240
 QY 99 LeuGlnIleAsnAsnLeuValSerGlyLeuAspMetAlaThrTyrPheCysAlaArgTyrPhe 118
 Db 241 TTACGATCAACGCTCAGAGCGGCTGATGATTTCTGTACAAAG---CGG 297

Db	658	TATTGTCAACGATGATGAGTCTCCGTCGACGTTGCGTGGAGGACCACTTGAGATG	717
Qy	259	Lys	259
Db	718	AAA	720

Search completed: April 20, 2005, 17:07:32
Job time : 812.902 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2005, 06:53:52 ; Search time 17.8485 Seconds
(without alignments)
2485.142 Million cell updates/sec

Title: US-09-596-774-6
Perfect score: 2441
Sequence: 1 MAWWTLLFLMAAKVPRQI.....LSTATKDTYDALHMOTLAPR 461

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_79:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	713	29.2	164	2 A40104	T-cell receptor CD
2	694	28.4	268	2 A56446	Ig heavy chain V r
3	665.5	27.3	249	2 S41374	single chain Fv an
4	611.5	25.1	233	2 JC5322	p53 specific singl
5	605	24.8	177	2 S54817	T-cell receptor lo
6	605	24.8	206	2 A35900	T-cell receptor CD
7	604.5	24.8	163	2 A31768	T-cell receptor ze
8	589.5	24.1	188	2 A45089	T-cell receptor CD
9	588.5	24.1	136	2 S35759	BHD9D10 protein -
10	585	24.0	146	4 S33905	Ig heavy chain pre
11	555	22.7	139	2 PH1225	Ig heavy chain pre
12	539	22.1	166	2 JC4654	T-cell receptor ze
13	539	22.1	166	2 T46424	T-cell surface gly
14	537.5	22.0	120	2 B42848	L6 mab heavy chain
15	536	22.0	124	2 PH1404	Ig heavy chain V r
16	522.5	21.4	118	2 S37204	Ig heavy chain V r
17	519	21.3	119	2 A53285	Ig heavy chain V a
18	516	21.1	119	2 H45722	anti-glycoprotein
19	510	20.9	117	2 S32190	Ig heavy chain V r
20	499.5	20.5	115	2 S19965	Ig heavy chain V r
21	495	20.3	107	2 S32192	Ig kappa chain V r
22	490	20.1	107	2 S32191	Ig kappa chain V r
23	490	20.1	113	2 B36259	Ig heavy chain V r
24	488.5	20.0	118	2 A32530	Ig heavy chain V r
25	486.5	19.9	115	2 S19968	Ig heavy chain V r
26	486	19.9	117	2 S32187	Ig heavy chain V r
27	486	19.9	118	2 S19967	Ig heavy chain V r
28	486	19.9	149	1 KWS511	Ig kappa chain pre
29	484.5	19.8	109	2 S26325	Ig heavy chain V r

30	484.5	19.8	120	2 S19963	Ig heavy chain V r
31	484	19.8	119	2 B32530	Ig heavy chain V r
32	481.5	19.7	134	2 S21916	Ig heavy chain V r
33	480	19.7	119	2 P00265	Ig kappa chain V r
34	477.5	19.6	114	2 D32967	Ig heavy chain V r
35	472.5	19.4	114	2 C32967	Ig heavy chain V r
36	470.5	19.3	142	2 S19245	Ig heavy chain pre
37	469.5	19.2	105	2 S24765	Ig heavy chain V r
38	469	19.2	152	2 S30751	Ig heavy chain pre
39	460	18.8	114	2 P10256	Ig heavy chain V r
40	458.5	18.8	105	2 S24766	Ig heavy chain V r
41	457.5	18.7	105	2 S24764	Ig heavy chain V r
42	453.5	18.6	108	2 P10083	Ig kappa chain V r
43	452	18.5	117	2 S42466	Ig kappa chain V r
44	451.5	18.5	105	2 S24763	Ig heavy chain V r
45	451	18.5	131	2 S26792	Ig heavy chain V r

ALIGNMENTS

RESULT 1
A40104
T-cell receptor CD3 zeta chain - mouse
C:Species: Mus musculus (house mouse)
C>Date: 20-Mar-1992 #sequence_revision 20-Mar-1992 #text_change 09-Jul-2004
C:Accession: A40104; I55293
R:Weisman, A.M.; Baniyash, M.; Hou, D.; Samuelson, L.E.; Burgess, W.H.; Klausner, R.D.
Science 239, 1018-1021, 1988
A:Title: Molecular cloning of the zeta chain of the T cell antigen receptor.
A:Reference number: A40104; MUID:88145643; PMID:3278377
A:Accession: A40104
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-164 <WEB>
A:Cross-references: UNIPROT:P24161; GB:M19729; NID:G201131; PID:AAA40171.1; PID:G201132
R:Baniyash, M.; Hsu, V.W.; Seldin, M.F.; Klausner, R.D.
J. Biol. Chem. 264, 13252-13257, 1989
A:Title: The isolation and characterization of the murine T cell antigen receptor zeta c
A:Reference number: I55293; MUID:88327299; PMID:2787796
A:Accession: I55293
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-164 <RES>
A:Cross-references: GB:J04967; NID:G556326; PID:AAA50301.1; PID:G556327
C:Gene: Tcrz
C:Keywords: phosphoprotein; T-cell receptor; transmembrane protein
Query Match 29.2%, Score 713; DB 2; Length 164;
Best local similarity 99.3%; Pred. No. 4e-42; Mismatches 1; Indels 0; Gaps 0;
Matches 138; Conservative 0;
OY 323 LEDPKLCYLDGILFIYGVITITLALYLRAKFSRSASMTANLQDPNOLYNELNGRREEDV 382
DB LLDRLKCLLDGILFIYGVITITLALYLRAKFSRSASMTANLQDPNOLYNELNGRREEDV 85
OY 383 LEKRRADPEWGGKQQRNRNPQEGVYNALQDKMAEAYSEIGTGERRRGKGDLYGGL 442
DB LEKRRADPEWGGKQQRNRNPQEGVYNALQDKMAEAYSEIGTGERRRGKGDLYGGL 145
OY 443 STATKDTYDALHMOTLAPR 461
DB STATKDTYDALHMOTLAPR 164
RESULT 2
A56446
Ig heavy chain V region (3H-3H scFv) - mouse (strain BALB/C)
C:Species: Mus musculus (house mouse)
C>Date: 19-Jan-1996 #sequence_revision 19-Jan-1996 #text_change 16-Aug-1996
C:Accession: A56446
R:Tang, P.M.; Foltz, L.A.; Mahoney, W.C.; Schueler, P.A.

J. Biol. Chem 270, 7829-7835, 1995
 A:Title: A high affinity digoxin-binding protein displayed on M13 is functionally identical
 A:Reference number: A56446; MUID:95229583; PMID:771873
 A:Accession: A56446
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-268 <TAN>
 A:Cross-references: CB:U20617
 C:Keywords: heterodimer; Immunoglobulin

Query Match	28.4%	Score 694	DB 2	Length 268
Best Local Similarity	55.2%	Pred. No. 1.5e-40		
Matches 133	Conservative 43	Mismatches 63	Indels 2	Gaps 2

```

QY  QIOLVSGPELKKPGETTVKISCKASGYPTNTYGMNMMVQAOAGGQIKMMGMITNSTGESTF 78
D  :::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY  QVQLQSSGAEIVKPKASVYKLTSTSTSGFNKIDITYHMVQKREQGLEIMIGRIAPANGITTKY 62
D  3

QY  ADDPKGRFDFSELTSTANTAYIQLINLKSEDMATYFCARMEVHYGVPMYGGTTVTVSSG 138
D  79

QY  DPKFOKATITADTSSNTAYIQLTSSLTSEDTLVVYICASYFLTR-YENVMGGTTVTVSSG 121
D  63

QY  GGGSGGGSGCGGSGSDIQLTQSHKFLSTVGRPVSTICKASQDYVNAYNAWYQKRGSGPKL 198
D  139

QY  GGGSGGGSDGCGGSDIELTQSPALMSASLGEKATVMSCASSV-NFIYTWQKSDASPKL 180
D  122

QY  LIYSASSRYTVGPSRFTSGSGSPDFTTITSSVOAEDLAVVPCQGHFRPTPTFGSGTKLEI 258
D  199

QY  WYVYITSHLPVGPVAPRFGSGSGNSYSYSLTSSWEGEDATVYCCQFTSPPTFGSGTKLEI 240
D  181

QY  K 259
D  259

QY  K 241
D  241

```

RESULT 3

single chain Fv antibody - mouse
C/Species: Mus musculus (house mouse)
C/Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 06-Jan-1995
C/Accession: S41374
R/Artseenko, O.; Weller, E.W.; Muentz, K.; Conrad, U.
submitted to the EMBL Data Library, January 1994
A/Description: Construction and functional characterization of a single chain Fv antibody
A/Reference number: S41374
A/Accession: S41374
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-249 <ART>
A/Cross-references: EMBL:Z29480

Query Match	27.3%;	Score 665.5;	DB 2;	Length 249;
Best Local Similarity	52.0%;	Pred. No. 1.2e-38;		
Matches 128;	Conservative 47;	Mismatches 66;	Indels 5;	Gaps 1;

```

QY      19 QIOLVOSGELKKPGETVIVISCACAGYPTIVYGNMWWAOAGOGGLKMMGIMNTSGSTF  78
      1 QVQLQSGAEIVIRPGASVYLTCTSHGCFNPKDITHWVQREKLEMTARIAPASGNVKY  60
Db
QY      79 ADDFKRFPSPISLETANNTIYLOINNLKSEDNATYFCARMEVYHGIVPYGGQTTVTVSSG  138
      61 VPRFQKAVITADTSSNTANVYLLLSLTSEBDNAVYVCARDRLTYLSLWGSGSTVTVSSR  120
Db
QY      139 GGGSGGGSGGGGGSDIOLTOSHKKFLSTVGDRVSLTCKASO-----DVYNAVAVYQCKPG  193
      121 GGGSGGGSGGGGGSDIELTQSPSPVYVIRPGESVTSICRSKSLLYSDODSYLFWFLRPG  180
Db
QY      194 GSPKLLIYASASRYGVSPRFTFGSGSGGDFEFTTISVYAEPLAVYFECQHRFTFTGSG  253
      181 GSPFOLLIYMSULASGVDPDRFSGSGSGISFTLRISVEALEVGVYIYCMQREYFLITGAG  240
Db
QY      254 TYLEIK 259
      ||||:|

```

Db 241 TKLELK 246

```

RESULT 4
JC5322
p53 specific single-chain antibody Pab421 - human
C:Species: Homo sapiens (man)
C:Date: 15-May-1997 #sequence_revision 15-May-1997 #text_change 18-Jul-1997
C:Accession: J05322

```

Query Match	25.1%;	Score 61.5;	DB 2;	Length 233;
Best Local Similarity	51.0%;	Pred. No. 5.8e-35;		
Matches 122;	Conservative 39;	Mismatches 67;	Indels 11;	Gaps 3;

QY 2 OSPELTKKGGEGYVYISCKRSGYYPFNNGYNNWKAQGGGLKMGWINTSTESFTADDFK 83
 Dd 2 ESABELVYRSGAVYKLSCTTSGFNINDYMHVKKRPGGLEMLGIDPENGDADMTBSG 61
 QY 84 GRDPSLETSANTAYLQIINNLSKEDMATYFCARMEVHYGYYPYMGCGTYYTVSSGGGSG 14
 Dd 62 VKRMIMADTSSNTAYLQISSLTSEDTAYYC-----NAGMDYMGCGTYYTVSSGGGSG 11

[illegible]

RESULT 5

T-cell receptor *iota* precursor - mouse
C/Species: Mus musculus (house mouse)
C/Date: 08-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 05-Nov-1999
C/Acession: S54817
R/Rocentini, G.; Ronchetti, S.; Baroli, A.; Testa, G.; d'Adamo, F.; Riccardi, C.; Migliorini, M.
Submitted to the EMBL Data Library, January 1995
A/Description: The T cell receptor *iota*: an alternatively spliced product of the T cell receptor gene
A/Reference number: S54817
A/Acession: S54817
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-177 <NC>
A/Cross-references: EMBL:X84237, NID:g809045, PIN:CAAS9015.1, PID:g809046
Keywords: T-cell receptor

Query Match	24.8%	Score 605	DB 2	Length 177
Best Local Similarity	99.2%	Pred. No. 1.2e-34		
Matches 117; Conservative	0	Mismatches 1	Indels 0	Gaps 0

QY 333 LEDPLCYLIDGLITFIYGVITITATYIPAKSSBSAETANLQOPNOLYNLNGREEBYD 382

Db 26 LIDPLCYLIDGLITFIYGVITITATYIPAKSSBSAETANLQOPNOLYNLNGREEBYD 85

QY 393 LEKTPARDPEMGKQOORRNPOEGVYNALQDMAEYSEIGTKGERRRGKHDGIYQ 440

86 LEKTPARPENMGKQOORRNPOEGVYNALQDMAEYSEIGTKGERRRGKHDGIYQ 143

RESULT 6
A35900
T-cell receptor CD3 eta chain precursor - mouse

C:Species: Mus musculus (house mouse)
C:Date: 23-Oct-1990 #sequence_revision 23-Oct-1990 #text_change 09-Jul-2004
C:Accession: A35900; #Accession: A60374; #Accession: A65522
R:Jin, Y.J.; Clayton, L.K.; Howard, F.D.; Koyasu, S.; Steh, M.; Steidrich, R.; Tarr, G.
Proc. Natl. Acad. Sci. U.S.A. 87, 3319-3323, 1990
A:Title: Molecular cloning of the CD3epsilon subunit identifies a CD3epsilon-related product in
A:Reference number: A35900; MUID:90239005; PMID:2139725
A:Accession: A35900
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-206 <JIN>
A:Cross-references: UNIPROT:P29020; GB:M3158; NID:G192488; PIDN:AAA37398.1; PID:G309159
R:Ohno, H.; Saito, T.
Int. Immunol. 2, 1117-1119, 1990
A:Title: CD3zeta and eta chains are produced by alternative splicing from a common gene.
A:Reference number: A60374; MUID:91190781; PMID:2150596
A:Accession: A60374
A:Status: not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 144-206 <OHNO>
R:Jensen, J.P.; Cenciarelli, C.; Hou, D.; Rellahan, B.L.; Dean, M.; Weissman, A.M.
J. Immunol. 150, 122-130, 1993
A:Title: T cell antigen receptor-eta subunit. Low levels of expression and limited cross
A:Reference number: A46522; MUID:93107707; PMID:8417118
A:Contents: annotation
C:Comment: The functional significance of this alternatively spliced product of the CD3
epsilon acids, differs widely among various mammalian species in sequence, length, and even
C:Keywords: alternative splicing; T-cell receptor; transmembrane protein

Query Match 24.8%; Score 605; DB 2; Length 206;
Best Local Similarity 99.2%; Pred. No. 1.4e-34;
Matches 117; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 323 LEDPRLCYLIDGLIFIVITITALLYLRKFSRSASATAANLDPNOLYNELNGRREEDV 382
DB 26 LLDPLCYLIDGLIFIVITITALLYLRKFSRSASATAANLDPNOLYNELNGRREEDV 85

OY 383 LEKRRADPEMGKQORRRNPOEGVYNALQKDKMAEAYSEIGTGERRRGKHGDLQ 440
DB 86 LEKRRADPEMGKQORRRNPOEGVYNALQKDKMAEAYSEIGTGERRRGKHGDLQ 143

RESULT 7
A31768
T-cell receptor zeta chain precursor - human
C:Species: Homo sapiens (man)
C:Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 09-Jul-2004
C:Accession: A31768
R:Weissman, A.M.; Hou, D.; Orloff, D.G.; Modi, W.S.; Seunarez, H.; O'Brien, S.J.; Klausne
Proc. Natl. Acad. Sci. U.S.A. 85, 9709-9713, 1988
A:Title: Molecular cloning and chromosomal localization of the human T-cell receptor zeta
A:Reference number: A31768; MUID:89071765; PMID:2974162
A:Accession: A31768
A:Molecule type: mRNA
A:Residues: 1-163 <WEI>
A:Cross-references: UNIPROT:P20963; GB:J04132; NID:G623041; PIDN:AAA60394.1; PID:G623042
C:Keywords: phosphoprotein; T-cell receptor; transmembrane protein
F:1-21/Domain: signal sequence #status predicted <SIG>
F:22-163/Product: T-cell receptor zeta chain #status predicted <MAT>

Query Match 24.8%; Score 604.5; DB 2; Length 163;
Best Local Similarity 84.9%; Pred. No. 1.2e-34;
Matches 118; Conservative 6; Mismatches 14; Indels 1; Gaps 1;

OY 323 LEDPRLCYLIDGLIFIVITITALLYLRKFSRSASATAANLDPNOLYNELNGRREEDV 382
DB 26 LLDPLCYLIDGLIFIVITITALLYLRKFSRSASATAANLDPNOLYNELNGRREEDV 85

OY 383 LEKRRADPEMGKQORRRNPOEGVYNALQKDKMAEAYSEIGTGERRRGKHGDLQ 442
DB 86 LDKRRGRDPEMGK-QRRKNPOEGVYNELQKDKMAEAYSEIGTGERRRGKHGDLQ 144

OY 443 STATKDTYDALHMQALPR 461

DB 145 STATKDTYDALHMQALPR 163
|||||
RESULT 8
A45089
T-cell receptor CD3 theta chain, alternate splice form - mouse
C:Species: Mus musculus (house mouse)
C:Date: 30-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
C:Accession: A45089; #Accession: A45089
R:Clayton, L.K.; Diener, A.C.; Lerner, A.; Tse, A.G.; Koyasu, S.; Reinherz, E.L.
J. Biol. Chem. 267, 26023-26030, 1992
A:Title: Differential regulation of T-cell receptor processing and surface expression afi
A:Reference number: A45089; MUID:93100325; PMID:1464613
A:Accession: A45089
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-188 <CLAY>
A:Cross-references: GB:S51932; NID:G261990; PIDN:AAB24559.1; PID:G261999
A:Experimental source: thymus
R:Lerner, A.; Diener, A.C.; Reinherz, E.L.; Clayton, L.K.
Eur. J. Immunol. 22, 2135-2140, 1992
A:Title: Human genomic sequences corresponding to murine CD3epsilon-related transcripts: lacc
A:Reference number: A49587; MUID:92347411; PMID:1322304
A:Accession: A49587
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-188 <RES>
A:Cross-references: GB:L03353; NID:G192508; PIDN:AAA37401.1; PID:G192509
C:Keywords: T-cell receptor

Query Match 24.1%; Score 589.5; DB 2; Length 188;
Best Local Similarity 98.3%; Pred. No. 1.5e-33;
Matches 116; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

OY 323 LEDPRLCYLIDGLIFIVITITALLYLRKFSRSASATAANLDPNOLYNELNGRREEDV 382
DB 26 LLDPLCYLIDGLIFIVITITALLYLRKFSRSASATAANLDPNOLYNELNGRREEDV 85

OY 383 LEKRRADPEMGKQORRRNPOEGVYNALQKDKMAEAYSEIGTGERRRGKHGDLQ 440
DB 86 LEKRRADPEMGK-QRRKNPOEGVYNALQKDKMAEAYSEIGTGERRRGKHGDLQ 142

RESULT 9
S35759
BHD9D10 protein - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 21-Jan-2000
C:Accession: S35759
R:Froyen, G.F.V.
submitted to the EMBL Data Library, May 1993
A:Reference number: S35759
A:Accession: S35759
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-136 <FRO>
A:Cross-references: EMBL:X72796; NID:G312496; PIDN:CAAS1316.1; PID:G312497
C:Superfamily: immunoglobulin V region; immunoglobulin homology
F:34-117/Domain: immunoglobulin homology <IMM>

Query Match 24.1%; Score 588.5; DB 2; Length 136;
Best Local Similarity 82.6%; Pred. No. 1.2e-33;
Matches 114; Conservative 7; Mismatches 14; Indels 3; Gaps 2;

OY 1 MAWVTLFLTAAAK-VPKQIQVQSGPELKKPGETVKSCKASGYPFTYGMWVQAP 59
DB 1 MAWVTLFLTAAQSAQAOQVQSGPELKKPGETVKSCKASGYFTYGMWVQAP 60

OY 60 GQGLKMGWINTSTGESSTPADDFGRDFSLTSSANTAYVQINNLKEDMATYFCARWEV 119
DB 61 GQGLKMGWINTYGESSTYVDDFKGRFVFSLETSSASAAVQINNLKEDMATYFCARGF 120

```

QY      120 YHGVPYWGQGTIVTVSS 137
      | : |||||: |||||
Db      121 Y--AMDYWGQGTSTVTVSS 136

```

RESULT 10
S33905
Ig heavy chain precursor V region - sythetic
C/Species: sythetic
C/Date: 13-Jan-1995 #sequence_revision 30-Apr-1998 #text_change 20-Oct-2000
C/Accession: S33905
R/Liu, A.Y.; Robinson, R.R.; Hellstroem, K.E.; Murray Jr., E.D.; Chang, C.P.; Hellstroem
Proc. Natl. Acad. Sci. U.S.A. 84, 3433-3443, 1987
A/Title: Chimeric mouse-human IgG1 antibody that can mediate lysys of cancer cells.
A/Reference number: S33905, MUID:87204152, PMID:3106970
A/Accession: S33905
A/Molecule type: mRNA
A/Residues: 1-146 <R1U>
A/Cross-references: EMBL:M16072; NID:G195270; PIDD:AAA38229.1; PID:G195271

Query Match	24.0%;	Score 585;	DB 4;	Length 146;
Best Local Similarity	80.6%;	Pred. No. 2.2e-33;		
Matches 112; Conservative	9;	Mismatches 16;	Indels 2;	Gaps 2

[illegible]

RESULT 11
PH1225
Ig heavy chain precursor V region (M-T151) - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 21-Jan-2000
C:Accession: PH1225
R:Reisemörm, W.; Scheuer, W.; Kaluza, B.; Schwirzke, M.; Reiter, C.; Flieger, D.; Lenz
Gene 121, 2771-278, 1992
A:Title: Combinatorial functions of two chimeric antibodies directed to human CD4 and or
A:Reference number: PH1224; MUID:93077041; PMID:1446824
A:Accession: PH1225
A:Molecule type: mRNA
A:Residues: 1-139 <WEI>
A:Cross-references: GB:550263; NID:9260763; PIN:AA82319.1; PID:9260764
A:Note: this mouse sequence was hybridized and fused with a human constant region gene
C:Keywords: heterotrimer; immunoglobulin V region; immunoglobulin homology
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-139/Product: Ig heavy chain V region #status predicted <MAT>
F:34-117/Domain: immunoglobulin homology <IMM>

Query Match	22.7%;	Score 555;	DB 2;	Length 139;
Best Local Similarity	76.3%;	Pred. No. 2.4e-31;		
Matches 106;	Conservative 13;	Mismatches 18;	Indels 2;	Gaps 2

```
QY      1 MAMVWTLTFLFMAAAR-VPRKQIDLVOSGELKKPGFTVAIXISKAGGYPTFYNNMVRQAP  59
          : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db      1 MAMVWTLTFLFMAAASIQAQIQLVOSGELTKTPGTVAISCKAGGYPTFDYSIHNVQAP  60
          : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
QY      60 GGGKMKMGINTSTEBSTFADDFEKGKRFPSLETSAHTAYLIQINNLSKEDMATVFCA-RWE  118
          : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
Db      61 GDLKLMMGIINTEGETPYADDFTGRFAFSLETSASTYLQINNLSKNEEDSTYCALHY  120
          : | : | : | : | : | : | : | : | : | : | : | : | : | : | : | :
QY      119 VYHGIVPVTGGGCTTVSSS  137
          | : | : | : | : | : |
```

Db 121 AYGDPLDYWGQTSVTSS 139

RESULT 12
JC4664
T-cell receptor zeta chain - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 10-May-1996 #sequence_revision 19-Jul-1996 #ext_change 05-Nov-1999
C:Accession: JC4664
R:Hagen, G.; Galle, Y.; Glaser, I.; Davis, W.C.; Baldwin, C.L.; Clevers, H.; Dobbelaert
Gene 169, 165-171, 1996
A:Title: Cloning, sequencing and expression of the bovine CD3 epsilon and TCR-zeta chains
A:Reference number: JC4663; PMID:96194796; PMID:864741
A:Accession: JC4664
A:Molecule type: mRNA
A:Residues: 1-166 <HAG>
A:Cross-references: GS:U25688; NID:G1263011; PIDN:AAC48548.1; PID:G1263012
A:Comment: This protein plays a pivotal role in linking T-cell receptor-triggering to se
lymphokine receptor gene expression.
C:Genetics:
A:Gene: tcr-zeta
C:Keywords: GTP binding; signal transduction; T-cell receptor
;129-146/Region: GDP/GTP-binding

Query Match	22.1%;	Score 539;	DB 2;	Length 166;
Best Local Similarity	75.2%;	Pred. No. 3.7e-30;		
Matches 106; Conservative	11;	Mismatches 22;	Indels 2;	Gaps 1.

QY 323 LDDPKCYLLDDILFYIGVITATLALTRKFSRSSETAMNLODPMLQVNEINLGRREYDV 382
 Db 26 LDDPKCYLLDDILFYIGVITATLALTRKFSRSANAPAYCOGQFPYNEINLVGRREYAV 85
 QY 383 LKKKARADPEWGCKOORRRNPOEGEYVNAJOKDKMAEAYSEIGTC--GERRRGKHGLTYO 440
 Db 86 LDRRGGFDPKGGKQRKKKNPEYVYNEJLRDKDAEAYSEIGMKSNDQRRRGKHDSGLTYO 145
 QY 441 GLSTATKDTYDALHMOTLAPR 461
 Db 146 GLSTATKDTYDALHMOTALPR 166

RESULT 13
I46424 T-cell surface glycoprotein CD3 zeta chain - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C:Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 09-Jul-2004
C:Accession: I46424, S22980
R:Hein, W.R.; Tunncliffe, A.
Immunogenetics 37, 279-284, 1993
A:Title: Invariant components of the sheep T-cell antigen receptor: cloning of the CD3 epsilon chain
A:Reference number: I46424; MUID:9313105; PMID:8420837
A:Accession: I46424
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-166 <HEI>
A:Cross-references: UNIPROT:P29329, EMBL:Z12968, NID:g13399, PIDD:CAA78312.1, PID:g1400
Keywords: glycoprotein

Query Match	22.1%;	Score 539;	DB 2;	Length 166;
Best Local Similarity	74.5%;	Pred. No. 3.7e-30;		
Matches 105; Conservative	13;	Mismatches 21;	Indels 2;	Gaps 1;

QY	32	LEDDPGLCYLLGGIIIF	YGVIIITALLYLAKF	FSRSAMETAN	LODPQLYNENL	NGRREYDV	362
QY	33	LEDDPGLCYLLGGIIIF	YGVIIITALLYLAKF	FSRSAMETAN	LODPQLYNENL	NGRREYDV	362
Db	26	LDDPGLCYLLGGIIIF	YGVIIITALLYLAKF	FSRSAMETAN	LODPQLYNENL	NGRREYDV	85
QY	383	LEKKARAPBEMGKQOAPRN	POEYVNALQD	KMAEYSEIGT	-GERRRGCHOST	LYQ	440
QY	39	LEKKARAPBEMGKQOAPRN	POEYVNALQD	KMAEYSEIGT	-GERRRGCHOST	LYQ	440
Db	86	LDRRGFPBEMGKQOAPRN	POEYVNALQD	KMAEYSEIGT	-GERRRGCHOST	LYQ	145
QY	441	GLSTATKDTYDALHMO	TALAPR	461			
QY	146	GLSTATKDTYDALHMO	TALAPR	166			
Db	146	GLSTATKDTYDALHMO	TALAPR	166			

RESULT 14

B42848
I6 mAb heavy chain V region - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 27-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jan-2000
C:Accession: B42848; S33903
R:Feil, H.P.; Gayle, M.A.; Yelton, D.; Lipsich, L.; Schieven, G.L.; Marken, J.S.; Aruffo
J. Biol. Chem. 267, 15552-15558, 1992
A:Title: Chimeric I6 anti-tumor antibody. Genomic construction, expression, and character
A:Reference number: A42848; MUID:92348410; PMID:1639794
A:Accession: B42848
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-120 <FE1>
A:Cross-references: GB:M90690; NID:g195065; PUID:AAA8146.1; PID:g195066
A>Note: Sequence extracted from NCBI backbone (NCBIN:109960, NCBI:P:109961)
A:Accession: S33903
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-120 <FE2>
A:Cross-references: EMBL:M90691
C:Superfamily: Immunoglobulin V region; immunoglobulin homology
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 22.0%; Score 537.5; DB 2; Length 120;

Best Local Similarity 84.2%; Pred. No. 3.2e-30;

Matches 101; Conservative 7; Mismatches 11; Indels 1; Gaps 1;

OY 19 QIQVQSGPELKKPGETVTKISCKASGYPTNYGMNWKQAPQGLKMMGWINTSTGESTF 78
DB 1 QIQVQSGPELKKPGETVTKISCKASGYPTNYGMNWKQAPQGLKMMGWINTSTGESTF 60
OY 79 ADDRGRDFSELSANTAYLQINNLSKEDMATYFCARWEVYHG-YVPYWGQSTIVTVSS 137
DB 61 ADDRGRDFSELSANTAYLQINNLSKEDMATYFCARWEVYHG-YVPYWGQSTIVTVSS 120

RESULT 15

PH1404
Ig heavy chain V region (clone micro m+ 46-12) - mouse (fragment)

C:Species: Mus musculus (house mouse)

C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 21-Jan-2000

C:Accession: PH1404; PH1406

R:Shirasawa, T.; Miyazoe, I.; Higashimura, S.; Kimoto, H.; Shigemoto, K.; Taniguchi, M.; Ta
J. Exp. Med. 176, 1209-1214, 1992

A:Title: Heavy chain variable (VH) region diversity generated by VH gene replacement in
a virus.
A:Reference number: PH1403; MUID:93018837; PMID:1402663

A:Accession: PH1404

A:Molecule type: DNA

A:Residues: 1-124 <SH1>

A:Accession: PH1406

A:Molecule type: DNA

A:Residues: 115-121 <SH2>

C:Genetics:

A:Insertions: 16/1

C:Superfamily: immunoglobulin V region; immunoglobulin homology

C:Keywords: heterotetramer; immunoglobulin

F:35-118/Domain: immunoglobulin homology <IMM>

Query Match 22.0%; Score 536; DB 2; Length 124;

Best Local Similarity 85.6%; Pred. No. 4.2e-30;

Matches 101; Conservative 6; Mismatches 9; Indels 2; Gaps 1;

OY 1 MAMVTLFLMAAAK--VPKQIQVQSGPELKKPGETVTKISCKASGYPTNYGMNWKQ 58
DB 1 MDMLNLLFLMAAAQTGAQIQVQSGPELKKPGETVTKISCKASGYPTNYGMNWKQ 60
OY 59 PQQGLKMMGWINTSTGESTFADDFKGRDFSELSANTAYLQINNLSKEDMATYFCAR 116
DB 61 PQQGLKMMGWINTSTGESTFADDFKGRDFSELSANTAYLQINNLSKEDMATYFCAR 118

Search completed: April 20, 2005, 07:07:48
Job time : 18.8485 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 20, 2005, 10:13:45 ; Search time 4523.31 Seconds
(without alignments)
3879.380 Million cell updates/sec

Title: US-09-596-774-6
2441
Perfect score: 1 MAWVTLFLMAAKVPRQI.....LSTATKTYDALHMQTLAPR 461
Sequence:

Scoring table:
BLASTSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlh
-Q=/cgm2_1/USPTO.spool_h/US09596774/runat.20042005.075443.12521/app.query.fasta_1.1230
-DB=EST -QMT=fastcap -SUFFIX=apr20.rst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bite -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -XGAPOP=10 -XGAPEXT=0
-USER=US09596774.OCGN.1.1.4591.0runat.20042005.075443.12521 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERRY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database :

EST:
1: gb_est1.*
2: gb_est2.*
3: gb_hlc.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_g981.*
9: gb_g982.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	749.5	30.7	671	AF240168	Mus muscu
2	715	29.3	953	BY746048	Mus muscu
3	715	29.3	1570	AK088087	Mus muscu
4	713	29.2	577	CA577711	KO713C05
5	704.5	28.9	1598	AK017904	Mus muscu
6	697.5	28.6	492	AY403844	Mus muscu
7	681.5	27.9	764	CO568334	AGENCOURT
8	681.5	27.9	786	CO572391	AGENCOURT
9	628	25.7	993	BI832527	603082128

c	10	625	25.6	872	7	CK629396	AM0-AA001
	11	620	25.4	618	7	CV029513	8329 Full1
	12	620	25.4	772	4	BI838213	603083131
	13	620	25.4	1038	5	BW921412	BM921412
	14	620	25.4	1113	1	AL554350	AL554350
	15	620	25.4	1546	3	CR625818	CR625818
	16	620	25.3	1557	3	CR601423	CR601423
	17	617.5	25.3	547	5	BQ562923	Full-1eng
	18	615	25.2	587	5	BP367042	BP367042
	19	614	25.2	1025	1	AL558350	AL558350
	20	607.5	24.9	899	7	CO648852	ILUMIGEN
	21	606.5	24.8	601	5	BQ474958	carabub49
	22	604.5	24.8	492	9	AY403842	Home_sapi
	23	604.5	24.8	575	5	BP367299	BP367299
	24	604.5	24.8	581	5	BP368604	BP368604
	25	604.5	24.8	712	1	AL557555	AL557555
	26	604.5	24.8	875	5	BK464301	BK464301
	27	604.5	24.8	970	5	BQ051777	AGENCOURT
	28	604.5	24.8	1095	1	AL532484	AL532484
	29	604.5	24.8	1126	3	CR615998	CR615998
	30	602.5	24.7	1104	7	CK629846	AM2-AA002
c	31	596.5	24.4	580	5	BP367951	BP367951
	32	595	24.4	855	2	BE371942	BE371942
	33	578.5	23.7	581	5	BP366659	BP366659
	34	575.5	23.6	672	7	CK633068	AM3-AP001
	35	575.5	23.6	572	6	CA570877	CA570877
	36	572	23.4	522	6	CA570877	K0513A06
	37	567.5	23.2	852	4	BI819937	603035487
	38	565	23.1	675	5	BP163731	BP163731
	39	565	23.1	707	5	BP167815	BP167815
	40	565	23.1	727	5	BP164018	BP164018
	41	565	23.1	823	5	BP167326	BP167326
	42	565	23.1	835	5	BP163661	BP163661
	43	565	23.1	841	5	BP161017	BP161017
	44	565	23.1	841	5	BP164042	BP164042
	45	565	23.1	845	5	BP434480	BP434480

ALIGNMENTS

RESULT 1
AF240168
LOCUS Mus musculus MRP5 mRNA, partial cds.
DEFINITION AF240168
ACCESSION AF240168
VERSION AF240168.1 GI:13877288
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
HTC

REFERENCE
AUTHORS
TITLE
Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Eukaryota; Eutheria; Rodentia; Scleroglossa; Muridae; Murine; Mus.
1 (bases 1 to 671)
Cui,D.X., Zeng,G.Y., Wang,F., Xu,J.R., Ren,D.Q., Guo,Y.H.,
Tian,F.R., Yan,X.J., Hou,Y.,
Mechanism of exogenous nucleic acids and their precursors improving
the repair of intestinal epithelium after gamma-irradiation in mice
World J. Gastroenterol. 6 (5), 709-717 (2000)

JOURNAL
PUBMED
REFERENCE
AUTHORS
TITLE
2 (bases 1 to 671)
Cui,D., Zeng,G., Yan,X., Li,X. and Su,C.
Cloning of mouse genes related to repairing of intestinal
epithelium of the irradiated mice by treatment with the intestinal
RNA of mice of the same strain
Int. J. Radiat. Biol. Relat. Stud. Phys. Chem. Med. 19 (1), 71-80
(2001)

JOURNAL
REFERENCE
AUTHORS
TITLE
3 (bases 1 to 671)
Cui,D., Zeng,G., Yan,X., Wang,F., Tian,F., Ren,D., Zhao,T., Li,X.
and Su,C.
Direct Submission
Submitted (02-MAR-2000) Molecular Cloning Laboratory, Institute of
Genetic Diagnosis, Fourth Military Medical University, No. 17 Chang
Le West Road, Xi'an 710032, China

JOURNAL
FEATURES
Location/Qualifiers

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11(2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC Building Adenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

Source Location/Qualifiers

1..953

/organism="Mus musculus"

/mol_type="mRNA"

/strain="NOD"

/db_xref="taxon:10090"

/clone="E430003L03"

/issue_type="thymus"

/cell_type="thymic cells"

/clone_lib="RIKEN full-length enriched, 2 days neonate thymus thymic cells (NOD)"

ORIGIN

Alignment Scores:

Pred. No.: 1.28e-60 Length: 953

Score: 715.00 Matches: 146

Percent Similarity: 86.39% Conservative: 0

Best Local Similarity: 86.39% Mismatches: 11

Query Match: 29.29% Indels: 12

DB: 6 Gaps: 2

US-09-596-774-6 (1-461) x BY746048 (1-953)

293 SerProValHisProThrGlyThrsSerGlnProGlnArgProGluAspProGlu 312

94 TCGGCTGATCTCTCCACGTCGGCTTCCAGAGACG-----AGCAGCAGA 138

313 GlySerValIlybGlyThrGlyLeuAspPheLeuGluAspProGlyLeuGlyLeu 332

139 GCTTTGGCTCG-----CTGGATCCCAAACTGTGCTACTGCTA 177

333 AspGlyIleuPheIleTyGlyValIleIleThrAlaLeuTyLeuArgAlaIlybPhe 352

178 GATGGAATCTCTTCATCTACGAGTCATCATCACAGCCCTGTACTGAGAGCAAAATTC 237

353 SerTrpSerAlaIleuThrAlaAlaSerLeuGlnAspProGlnGlnIleuTyPheGlu 372

238 AGCAGAGAGTGAGAGACTGCTGCCAAGTGGAGGACCCCAACCAAGCTCTACAAATGAGCTC 297

373 AsnLeuGlyArgArgGluGluGlyTyPheValLeuGluIlybIlybArgAlaArgAspProGlu 392

298 AATCTAGGCGCAAGAGAGGAAATATCACTCTTGGAAGAAAGCGGCTCGGATTCACAG 357

393 MetGlyIlybGlnGlnArgArgArgPheProGlnGlnIlybValTyPheAlaIleuGln 412

358 ATGGAGGCAAAACAGCAGAGAGAGGAGAAACCCCGAAGCGGTATCAATGACACTGCGAG 417

413 LysAspIlybMetAlaGluAlaIleTySerGluIleGlyThrIlybGlyIlybValArgArgGly 432

418 AAAACAGAGATGGAGAAAGCTTACAGTGGACCAAAAAGGAGGAGGAGGAGAGG 477

433 LysGlyHisAspGlyLeuTyGlnIlybLeuSerThrAlaThrIlybAspThrTyPheAla 452

478 AAGGGGCGACGATGGCTTACAGAGGTCTTCAGCACTGCCACCAAGAACACTATGATGCC 537

453 LeuHisMetGlnThrLeuAlaProArg 461

538 CTGCATATGACAGACCTCGCCCTGCG 564

RESULT 3

AK088087

LOCUS

DEFINITION

Mus musculus 2 days neonate thymus thymic cells cDNA, RIKEN full-length enriched library, clone:E430003L03 product:CD3 antigen, zeta polypeptide, full insert sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AK088087 1570 bp. mRNA linear HTC 03-APR-2004

Mus musculus 2 days neonate thymus thymic cells cDNA, RIKEN full-length enriched library, clone:E430003L03 product:CD3 antigen, zeta polypeptide, full insert sequence.

AK088087

AK088087.1 GI:26104673

HTC: CAP trapper.

Mus musculus (house mouse)

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 Carninci, P. and Hayashizaki, Y.

High-efficiency full-length cDNA cloning

Meth. Enzymol. 303, 19-44 (1999)

99279253

10349636

2

Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.

Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes

Genome Res. 10 (10), 1617-1630 (2000)

20499374

11042159

3

Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitamura, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kasai, Y., Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matsubara, Y., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.

RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer

Genome Res. 10 (11), 1757-1771 (2000)

20530913

11076861

4

The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.

Functional annotation of a full-length mouse cDNA collection

Nature 409, 685-690 (2001)

5

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

6 (bases 1 to 1570)

Adachi, J., Aizawa, K., Akimura, T., Arikawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirokawa, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kigawa, I., Kusunaga, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, J., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Chemical Research, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: <http://genome.gsc.riken.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC Building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.

Please visit our web site for further details.
URL: <http://genome.gsc.riken.jp/>
URL: <http://fantom.gsc.riken.jp/>
Location/Qualifiers

FEATURES

source

1..1570
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NOD"
/db_xref="FANTOM DB:E430003L03"
/db_xref="taxon:10090"
/clone="E430003L03"
/cell_type="thymic cells"
/tissue_type="thymus"
/clone_id="RIKEN full-length enriched mouse cDNA library"
/dev_stage="2 days neonate"
76..571
/note="CD3 antigen, zeta polypeptide (MGD|MG:88334, GB|NM_031162, evidence: BLASTN, 100%, match=536), putative"
1549..1554
/note="putative"
1570
/note="putative"

ALIGNMENT SCORES:

Pred. No.: 2,6e-60 Length: 1570
Score: 715.00 Matches: 146
Percent Similarity: 86.39% Conservative: 0
Best Local Similarity: 86.39% Mismatches: 11
Query Match: 29.29% Indels: 12
DB: 3 Gaps: 2

US-09-596-774-6 (1-461) x AK088087 (1-1570)

233 SerProValHisProThrGlyThrSerGlnProGlnArgProGluAspGlyArgProArg 312
98 TCGCGATCGATCCTCCACGTCGCGGTTCCAGAGAGAG-----AGGCACAGA 142
313 GlySerValValGlyThrGlyLeuAspPheLeuGlnAspProGlyLeuGlyLeu 332
143 CTTTGGTCTCG-----CTGGATCCCAAACTCTGCTACTTGGCTA 181
333 AspGlyLeuPheLeuGlyThrGlyValValGlyLeuThrAlaLeuThrLeuAlaGlyPhe 352
182 GATGGAATCCTCTTCATCTACGGAGTCATATCACACCTCTGACCTGAGAGCAAAATTC 241
353 SerArgSerAlaGluThrAlaAlaAsnLeuGlnAspProAsnGlnLeuThrAsnGluLeu 372
242 AGCAGAGGTGAGAGACTGCTGCCCACTCCAGAGACCCCAACAGCTTCAATAGAGTTC 301
373 AsnLeuGlyArgGlnGluGlyThrAspValLeuGlnGlyLeuArgAlaArgAspProGlu 392
302 AACTTAGAGCGAAGAGAGAAATGACGTTCTTGAAGAAGAGGGGCTGGGATCCAGAG 361
393 MetGlyGlyValGlnGlnArgGlnArgGlnArgGlnArgGlnArgGlnArgGlnArgGln 412
362 ATGGAGGCGAAG 421
413 LysAspLysMetAlaGlnAlaThrSerGluValGlyThrGlyGlyLeuArgArgGly 432
422 AAAGACAAAGTGCACAAAGCTTCACTGATCGGACAAAGAGGAGAGAGAGAGAGAGAG 481
433 LysGlyHisAspGlyLeuThrGlnGlyLeuSerThrAlaThrLysAspThrLysAspAla 452

Db 482 AAGGGGACAGATGCGCTTTACAGGGTCTCAGCACTGCCACCAAGAGACACTATGATGCC 541
Qy 453 LeuHisMetGlnThrLeuAlaProArg 461
Db 542 CTGCATATGACAGACCTTGGCCCTCGC 568

RESULT 4
CA577711
LOCUS
DEFINITION
CA577711 577 bp mRNA linear EST 19-NOV-2002
CDNA library (Long) Mus musculus cDNA clone NIA:K0713C05
IMAGE:30074812 5', mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 577)
Piao, Y., Kargul, G.J., Dudekula, D.B., Qian, Y., Luo, A., Carter, M.G.,
Alba, K., Taub, D., Longo, D.L., Keller, J. and Ko, M.S.H.
Systematic Analyses of NIA Mouse Hematopoietic Stem Cell
(Lin-/C-Kit-/Sca-1-) cDNA library (Long)
Unpublished (2001)
Other ESTs: K0713C05-3
Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@igsun.grc.nia.nih.gov
Plate: K0713 row: C column: 05
Seq primer: M13 Reverse
High quality sequence stop: 577
POLYA=No.

FEATURES

source

1..577
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6NCR"
/db_xref="nlaEST:K0713C05-5N"
/db_xref="taxon:10090"
/clone="NIA:K0713C05 IMAGE:30074812"
/tissue_type="Hematopoietic Stem Cell
(Lin-/C-Kit-/Sca-1-)"
/dev_stage="Age approx. 10 weeks old"
/lab_host="DH10B"
/clone_id="NIA Mouse Hematopoietic Stem Cell
(Lin-/C-Kit-/Sca-1-) cDNA library (Long)"
/note="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2:
NotI; Mouse cDNA project by the Laboratory of Genetics,
National Institute on Aging (NIA), Intramural Research
Program, NIH (<http://igsun.grc.nia.nih.gov/cdna>). This is
a long-transcript enriched cDNA library (Ref. Genome Res.
11: 1553-1558 (2001)). [PMID: 11544199]. Total RNAs were
obtained from Drs. Dennis Taub, Dan Longo (National
Institute on Aging, USA), Jonathan Keller (National Cancer
Institute, USA). Double-stranded cDNAs were synthesized
with an Oligo(dT) primer (Invitrogen):
5'-pGACTAGTCTAGATGCGAGCGGCGGCTTTT-3' from
0.9 ug of total RNA, treated with T4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to Lome-Linker B-SalI, purified by phenol/chloroform, and
separated from free linkers by centrifugation 100. Then, the
cDNAs were amplified by long-range high fidelity PCR using
Ex Tag polymerase (Takara) with a primer SalI-S. The
products were purified by phenol/chloroform and centrifuged
100. The cDNAs were digested with SalI and NotI enzymes
and cloned into SalI/NotI site of pSPORT1 plasmid vector.
The DH10B E. coli host was transformed with the ligation
mixture by the standard chemical method. The average
insert size is about 2.1 kb. The library was constructed
by Yulan Piao (NIA)."

ORIGIN					
Alignment Scores:					
Pred. No.:	9,886-61	Length:	577		
Score:	713..00	Matches:	138		
Percent Similarity:	99.28%	Conservative:	0		
Best Local Similarity:	99.28%	Mismatches:	1		
Query Match:	29.21%	Indels:	0		
DB:	6	Gaps:	0		
US-09-596-774-6 (1-461) x CA577711 (1-577)					
Oy	323	LeugluAePProLyLeuCYeTYrLeuLeuaEepGIYileuPheIleTyGlyValille	342		
Dd	133	CTGCtGGATCCCAACTCTGCTACTTGTAGATGAATCCTTTATCTACGAGTCATC	192		
Oy	343	IleThralaleTYrLeuaArgAlalySpheSerARgSeraGlutThraAlaalaenleu	362		
Dd	193	ATCACAGCCCTGTACTCTGAGACCAAAATTACACGAGATGACAGACTGTGCCAACCTG	252		
Oy	363	GLNAPProAengInleuTYrAnqIUleuanleuGIYAgaRgIngluTYrApval	382		
Dd	253	CAGGACCCCAACCAGCTCTACCATATAGCTCAATCTAGGCGCAAGAGAATATACCTC	312		
Oy	383	LeugluYlBLYArAlARgAsPProGluwetGIYGIYLGInGLARARgARgAn	402		
Dd	313	TTCGAGAAAGACGGGCTCGGATCCAGATGGAGGCAACACACAGAGAGAGAHAC	372		
Oy	403	ProGIngluGIYValTYrAsnAlaleuGlnlyAsPlySMetalagUalATYSerglu	422		
Dd	373	CCCCAGGAAGCGCTATACATGSCATGCGAAGAACAGATGGCAGAAAGCTACAGTAG	432		
Oy	423	IleglyThrlYsglyGUarGARgArGlyLYSGlyHIsapGlyLeuTYrGInglyLeu	442		
Dd	433	ATCGGCAAAAAGCGGAGAGCGGAGAGGCAAGGCGCAAGATGGCTTTACAGAGGTC	492		
Oy	443	SerThrIAthrlyAsPThrTYrAspAlaleuHISmetGINThreulaAProkrg	461		
Dd	493	AGCACTGCCACCAAGACA CTTATATGCTCCCTGCATATGCAAGACCTTG3CCCTTGC	549		
RESULT 5					
AKO17904					
LOCUS					
AKO17904 1598 bp mRNA linear HTC 03-APR-2004					
DEFINITION Mus musculus adult male thymus cDNA, RIKEN full-length enriched library, clone:5830407D18 product:CD3 antigen, zeta polypeptide, full insert sequence.					
ACCESSION AKO17904					
VERSION AKO17904.1 GI:12857392					
KEYWORDS HTC; CAP trapper.					
SOURCE Mus musculus (house mouse)					
ORGANISM Mus musculus					
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.					
1 Carninci,P. and Hayashizaki,Y High-efficiency full-length cDNA cloning Mech. Enzymol. 303, 19-44 (1999)					
2 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Kono,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)					
JOURNAL MEDLINE PUBMED					
TITLE 20493374 11042159					
AUTHORS Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Kono,H., Akiyama,U., Nishi,K., Kitsuma,T., Tashiro,H., Itoh,M., Suni,N., Ishii,Y., Nakamura,S., Hazama,M., Nishine,T., Harada,A., Yamamoto,R., Matsunoto,H., Sakaguchi,S., Ikegami,T., Kasaiyagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Wataniki,M.,					
REFERENCE					

JOURNAL	Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.
REFERENCE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
GENOME	Genome Res. 10 (11), 1757-1771 (2000)
PMID	20530913
PMID	11076861
REFERENCE	4
AUTHORS	The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
TITLE	Functional annotation of a full-length mouse cDNA collection
REFERENCE	Nature 409, 685-690 (2001)
AUTHORS	5
TITLE	The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
REFERENCE	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
JOURNAL	Nature 420, 563-573 (2002)
REFERENCE	6 (bases 1 to 1598)
AUTHORS	Adachi, J., Aizawa, K., Akhira, S., Akimura, T., Arai, A., Aono, H., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hirotsu, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kikuchi, K., Kato, H., Kawai, J., Kojima, Y., Komuro, H., Kouda, M., Koyama, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, K., Saito, R., Sakai, C., Sakai, K., Sano, H., Saeki, D., Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T., Tajima, Y., Toyota, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
TITLE	Direct Submission
JOURNAL	Submitted (10-JUL-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Saitoh-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT	Please visit our web site (http://genome.gsc.riken.jp/) for further details.
FEATURES	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGGAGGATCGAGTATTAATTAATATCCGCCGCCGCC 3']. cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5'-GAGAGGAGGATCGAGTATTAATTAATATCCGCCGCCGCC 3']. cDNA was bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI, 3' end: BamHI. Host: DH10B.
FEATURES	Location/Qualifiers
SOURCE	1. 1598
FEATURES	/organism="Mus musculus"
SOURCE	/mol_type="mRNA"
FEATURES	/strain="C57BL/6J"
SOURCE	/db_xref="FANTOM,DB:5830407D18"
FEATURES	/db_xref="taxon:10090"
SOURCE	/clone="5830407D18"
FEATURES	/sex="male"
SOURCE	/tissue_type="thymus"
FEATURES	/clone_lib="RIKEN full-length enriched mouse cDNA library"
SOURCE	/dev_stage="adult"
FEATURES	112. 606
SOURCE	/note="unnamed protein product; CD3 antigen, zeta polypeptide (MGI:88334, GB NM_031162, evidence: BLASTN, 100%, match=536) putative"

ACCESSION IMAGE:7384006 5', mRNA sequence.
 VERSION COS68334
 KEYWORDS COS68334.1 GI:50380963
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 764)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics / NIH
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cga@dcf-remail.nih.gov
 Tissue Procurement: Drs. Josef Lazar & Howard Jacob, Medical
 College of Wisconsin
 CDNA Library Preparation: Open Biosystems
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 plate: LLM15548 row: d column: 20
 High quality sequence stop: 746.
 Location/Qualifiers
 1..764
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /clone="IMAGE:7384006"
 /lab_host="DH10B Tona"
 /clone_1db="NIH_MGC_250"
 /note="Organ: thymus; Vector: pExpress-1; Site: 1: EcoRV;
 Site 2: NotI; RNA obtained from testis tissue of 8 wk old
 animal. Tissues were snap-frozen and kept at -80C before
 RNA extraction and purification (TRI-reagent method). CDNA
 was primed using oligo-dT primer:
 5'-pGACTAGTCTGATGATCGGAGCGGCCCTTTA25-3' and cloned into
 the EcoRV/NotI sites of pExpress-1. Size-selection >1.25kb
 resulted in an average insert size of 1.9 kb. This is a
 primary library (normalized library is NIH_MGC_251) and
 was constructed by Open Biosystems. Note: this is a
 NIH_MGC library"

ORIGIN

Alignment Scores: 2.08e-57 Length: 764
 Pred. No.: 681.50 Matches: 140
 Score: 81.56% Conservative: 6
 Percent Similarity: 78.21% Mismatches: 21
 Best Local Similarity: 27.92% Indels: 12
 Query Match: 7 Gaps: 3

US-09-596-774-6 (1-461) x COS68334 (1-764)

QY 292 ProberProvalHisPro-----ThrgLYThrSerGlnPro 303
 Db 28 CTTGCTCTGGGCACTCTCCAGGCTTGCAAGATGAAGAGCGCATCGCTCGC 87
 QY 304 GlnArgProGluAapCyAargProArgGlySerValysGlyThrgLYLeu-AapPheLe 323
 Db 88 CTGATCCTTCAAGTGCAGTCCAGGAGC-----AGAGGCACAGAGCTTGGCTGCT 141
 QY 323 uGUAAPProLYLeuCYsTYrLeuLeuAapGlyLYLeuPheLeTYrGlyValIleI 343
 Db 142 G--GATCCCAACTGCTATATGCTAGAGGAACTCTTTCATCTACGAGATCATGCT 198
 QY 343 eThAlaLeuTYrLeuArgAlaLYsPheSerArgSerAlaGluThrAlaAaLeuG 363

Db 199 CACGACCCCTGTACTGAGAGCAAAATTCACAGAGAGTGCAGATCTGCTTACTTCA 258
 QY 363 iAaPProAaSnGlnLeuTYrAaSnGlnLeuAaSnLeuGLYAArgGlnGlnTYrAaPValLe 383
 Db 259 GGACCCCAACAGGCTCTTATACGAGCTCACTTGGGCGAAGAGAAATGTATGTTT 318
 QY 383 uGlnLYsAArgAlaAArgAaPProGluMeGlyLYsGlnGlnLYsAArgAArgAaPPr 403
 Db 319 GGACAGAGAGGCGCTCGGAGATCCGAGATGGGGGCAACAGCAGAGAGAGAAACCC 378
 QY 403 oGlnGlnGlyValTYrAaSnAlaLeuGlnLYsAArgLYsMeCaLaGlnAlaTYrSerGlnI 423
 Db 379 CCAGAAAGGCGCTGTACATGCACTGCAAGAAAGCAAGATGGCAGAGCTTACAGTGAAT 438
 QY 423 eGlyThrLYsGlyGlnLYsAArgAArgGlyLYsGlyLYsAArgLYsTYrGlnGlnLYsLe 443
 Db 439 TGGATGAAGAGCCAGAGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 498
 QY 443 rThAlaThrLYsAaPThrTYrAaPAlaLeuHisMeGlnThrLeuAlaProArg 461
 Db 499 CACTGCCACAGAGCACTATGACGCGCTGCATATGAGACCTGCCCCCTGCG 553

RESULT 8
 COS72391
 LOCUS COS72391
 DEFINITION AGENCOURT 28621196 NIH_MGC_250 Rattus norvegicus CDNA clone
 IMAGE:7381111 5', mRNA sequence.
 ACCESSION COS72391
 VERSION COS72391.1 GI:50385020
 KEYWORDS EST.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 786)
 NIH-MGC http://mgc.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics / NIH
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cga@dcf-remail.nih.gov
 Tissue Procurement: Drs. Josef Lazar & Howard Jacob, Medical
 College of Wisconsin
 CDNA Library Preparation: Open Biosystems
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 plate: LLM15540 row: 1 column: 05
 High quality sequence stop: 701.
 Location/Qualifiers
 1..786
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /clone="IMAGE:7381111"
 /lab_host="DH10B Tona"
 /clone_1db="NIH_MGC_250"
 /note="Organ: thymus; Vector: pExpress-1; Site: 1: EcoRV;
 Site 2: NotI; RNA obtained from testis tissue of 8 wk old
 animal. Tissues were snap-frozen and kept at -80C before
 RNA extraction and purification (TRI-reagent method). CDNA
 was primed using oligo-dT primer:
 5'-pGACTAGTCTGATGATCGGAGCGGCCCTTTA25-3' and cloned into
 the EcoRV/NotI sites of pExpress-1. Size-selection >1.25kb
 resulted in an average insert size of 1.9 kb. This is a
 primary library (normalized library is NIH_MGC_251) and
 was constructed by Open Biosystems. Note: this is a
 NIH_MGC library"

VERSION CK629396.1 GI:45753871
 KEYWORDS EST.
 SOURCE Apis mellifera (honey bee)
 ORGANISM Apis mellifera
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea; Apidae; Apis.
 1 (bases 1 to 872)
 REFERENCE Nunes,F.M.F., Valente,V., Sousa,J.F., Cunha,M.A.V., Pinheiro,D.G., Maia,R.M., Araujo,D.D., Costa,M.C.R., Martins,W.K., Carvalho,A.F., Moraes,N., Nascimento,A.M., Peixoto,P.M.V., Silva,M.F.R., Ramos,R.G.P., Reis,L.F.L., Dias Neto,E., Souza,S.J., Simpson,A.J.G., Zago,M.A., Soares,A.E.E., Bitondi,M.M.G., Espresafico,E.M., Espindola,F.S., Peco-Larson,M.L., Simoes,Z.L.P., Hartfelder,K. and Silva,W.A., Jr.
 TITLE Open reading frame ESTs - an efficient strategy for analysis of the honey bee transcriptome
 JOURNAL Unpublished (2004)
 COMMENT Contact: Silva Jr, W. A. Molecular Genetic and Bioinformatics Laboratory Department of Genetics, FMRP/USP, FUNDEP/Rua Tenente Catão Roxo, 2501, CEP 14051-140, Brazil
 Tel: +55 16 39639300
 Fax: +55 16 39639309
 Email: wilsonj@usp.br
 This sequence was derived from the FAPESP Genome Program High quality sequence start: 66
 High quality sequence stop: 625.
 FEATURES
 source
 1..872
 /organism="Apis mellifera"
 /mol_type="mRNA"
 /strain="Africanized"
 /db_xref="taxon:7460"
 /sex="female, worker"
 /dev_stage="adult"
 /clone_id="AA0013"
 /note="Organ: whole body"
 ORIGIN
 Alignment Scores:
 Pred. No.: 1,12e-51 Length: 872
 Score: 625.00 Matches: 127
 Percent Similarity: 66.14% Conservative: 41
 Best Local Similarity: 50.00% Mismatches: 76
 Query Match: 25.60% Gaps: 11
 DB: 7 Indels: 4
 US-09-596-774-6 (1-461) x CK629396 (1-872)
 QY 8 LeuPheLeuMetAlaAlaAlaValProLySgInIleGInLeuValGInSerGlyPro 27
 DB 846 ATGTTCTCTTACCGCACCGCCATGCC-----AGTGGACCG 808
 QY 28 GILeuLysLysProGlyGluThr---ValLysIleSerCysLysAlaSer---GlyTyr 45
 DB 807 CACCGTGGAAACGAATGTAACCTGGGCTTCATGAATTTCTCGCAAGCTTCTGCTAC 748
 QY 46 ProPheThrAnTYrGlyMetAsnTTPValysGInAlaProGlyGInGlyLeuLysTTP 65
 DB 747 ATCTTCACAGTTATGATATAGACTGGGTGAGCAGCGCTGAACAGGAGCTTGAGTGG 688
 QY 66 MetGlyTTPIleAsnThrSerThrGlyGluSerThrPheAlaAspPheLysGlyArg 85
 DB 687 ATTCGAGTGGTTTTCCTGGAGAGGGGAGTACTGAATTCATGAGAGTTCAAGGGCAGG 629
 QY 86 PheAspPheSerLeuGluThrSerAlaAsnThrAlaTyrLeuGInIleAsnLys 105
 DB 628 CCCACACTGAGTGGAGAGTCTCCACGACAGCTATATGAGCTCACTAGGCTGACA 569
 QY 106 SerGluAspMetAlaThrTyrPheCysAlaArgTTPGluValTyrHisGlyTyrValPro 125
 DB 568 TTCGAGAGCTGTGCTATTTCTGTGCTAGAGGGGACTACTATAGCGGCTACTTTGAC 509

QY 126 TyrTTPGlyGInGlyThrThrValThrValSerSerGlyGlyGlySerGlyGly 145
 DB 508 TTGTGGNCCAGGAGACACCGTACCGTCTCCATGATGGAGNCGGTTCAGGGGAGGT 449
 QY 146 GlySerGlyGlyGlyGlySerAspIleGInLeuThrGInSerHisLysPheLeuSerThr 165
 DB 448 GGCTCTGGCGGTGGCGATCTGATCATTTAGCTCACCCAGTCTCCAGCAATCATCTCTCA 389
 QY 166 SerValIleAspArgValSerIleThrCysLysAlaSerGInAspValTyrAsnAlaVal 185
 DB 388 TTCACAGGGAGAGAGGTCACCATACCTGACGTGCCAGCTTCAAGTATA---CGTTACATA 332
 QY 186 AlaTTPYrGInGInLysProGlyGInSerProLysLeuLeuIleTyrSerAlaSerSer 205
 DB 331 TATTGTATCCACAGCAAGACCTGATCTCTCCCGCAGACTCTCGATTATGACATCCAC 272
 QY 206 ArgTyrThrGlyValProSerArgPheThrGlySerGlySerGlyProAspPheThrPhe 225
 DB 271 GTGGCTCTCGAGTCCCTTTTCGCTTCAGTGGCAGTGGGTCTGGACCTTATTCTCTC 212
 QY 226 ThrIleSerSerValGInAlaGluAspLeuAlaValTyrPheCysGInGInHisPheArg 245
 DB 211 ACAATCAACCGAATGAGAGCTGAGATGCTGCCATTATTACTGCCAGAGTGAAGTGGT 152
 QY 246 ThrProPheThrPheGlySerGlyThrLysLeuGluLys 259
 DB 151 TATCCGTACAGTTCGAGGGGGGACCAAGCTGAGAGCTGAAA 110
 RESULT 11
 CVO29513 618 bp mRNA linear EST 20-AUG-2004
 LOCUS 8329 Pull Length cDNA from the Mammalian Gene Collection Homo
 DEFINITION sapiens cDNA 5' similar to BC025703, mRNA sequence.
 ACCESSION CVO29513
 VERSION CVO29513.1 GI:51487684
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 618)
 REFERENCE Rual,J.F., Hirozane-Kishikawa,T., Hao,T., Bertin,N., Li,S., Driscot,A., Li,N., Rosenbery,J., Lamesch,P., Vidalain,P.O., Clingingsmith,T.R., Hartley,J.L., Esposito,D., Cheo,D., Moore,T., Simmons,B., Sequerra,R., Bosak,S., Doucet-Stamm,L., Le Peuch,C., Vandenhaute,J., Quisick,M.E., Albala,J.S., Hill,D.E. and Vidal,M.
 TITLE Human ORFeome Version 1.1: a Platform for Reverse Proteomics
 JOURNAL Genome Res. (2004) In press
 COMMENT Contact: Vidal M
 Marc Vidal Laboratory
 Dana Farber Cancer Institute
 1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
 Tel: 617 632 5180
 Fax: 617 632 5739
 Email: Marc.Vidal@dfci.harvard.edu
 ORF Sequence Tag (OST) of Gateway Entry construct. Each cloned ORF
 template DNA and ORF specific primers
 PCR primers
 FORWARD: ATGAGTGAAGGCGCTTT
 BACKWARD: TAGCGAGGGGCGAGCGCTG
 Insert Length: 618 Std Error: 29.00
 Plate: 11027 Row: 09 Column: H
 Seq primer: ACTGGCCGTCGTTTACACAGCTGACGTGACGGAAGAAC
 High quality sequence start: 100
 High quality sequence stop: 617
 POLYA=NO.
 FEATURES
 source
 1..618
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"

/tissue_type="mixed"
/clone_lib="Full length cDNA from the Mammalian Gene
Collection"
/note="Vector: mixed. The ORFs were PCR amplified from the
MGC (Mammalian Gene Collection) as of April 2004 and
cloned by recombinational Gateway cloning into pDONR223
donor vector. Reference: MGC (Mammalian Gene Collection)
Program Team, Generation and Initial Analysis of more than
15,000 Full-length Human and Mouse cDNA Sequences. PNAS,
2002, 99(26), 16899-16903"

ORIGIN

Alignment Scores:

Pred. No.:	2,176-51	Length:	618
Score:	620.00	Matches:	126
Percent Similarity:	79.64%	Conservative:	7
Best Local Similarity:	75.45%	Mismatches:	21
Query Match:	25.40%	Indels:	13
DB:	7	Gaps:	2

US-09-596-774-6 (1-461) x CV029513 (1-618)

```

QY 296 HisProthrglyThrsSerGlnProGlnArgProGlnuSpCySargProArgGlySerVal 315
DB 30 CATCTCGAGGACAGATTGCCGATTACAGA-----59
QY 316 LysGlyThrGlyLeu-AspPheLeuGluAspProLysLeuGlySerThrLeuAspGlyI 335
DB 60 ---GGCACAGAGCTTTGGCTGCTG---GATCCCAAACTGCTACCTGCTGAGTGAAT 113
QY 335 eLeuPheIleTyrglyValIleIleThrAlaLeuTyrrLeuArgAlaLysPheSerArgse 355
DB 114 CTTCTTCATCATAGTGTGATCTTCTCAGTCTTGTCTGAGTGAAGTTCACAGAGAG 173
QY 355 ValAGluThrAlaAlaAsnLeuGlnAspProAsnGlnLeuTyrrAsnGluLeuAsnLeuG 375
DB 174 CGCAGACGCGCCCGCTACAGCCAGAGCGCCAGAACAGCTCATTAACGAGCTCAATCAGG 233
QY 375 YArgArgGluGluThrAspValLeuGluLysArgAlaArgAspProGluMetGlyI 395
DB 234 ACGAAGAGAGAGATGATGTTTGGACAAAGACGTGGCCGGGACCTGAGATGGGGG 293
QY 395 YLysGlnGlnArgArgArgAsnProGlnGlnGlyValIleTyrrAsnAlaLeuGlnLysAsp 415
DB 294 AAAGCCGAGAGAGAGAGAAACCTCAGAGAGGCTGTACATGAACGCAAGAAAGATA 353
QY 415 sMetAlaGlnAlaTyrrSerGlnIleGlyThrLysGlyGluArgArgGlyLysGlyH 435
DB 354 GATGGCGAGGCTTACAGTGAATTTGGGATGAAAGCGCGGAGGGGCAAGGGGCA 413
QY 435 sAspGlyLeuTyrrGlnGlyLeuSerThrAlaThrLysAspThrTyrrAspAlaLeuH 455
DB 414 CGATGGCTTTTACAGGCTCTCAGTACAGCACAAGGACCTTACGAGCCCTTCACAT 473
QY 455 tGlnThrLeuAlaProArg 461
DB 474 GCAAGCCTTGCCTCC 492

```

RESULT 12 772 bp mRNA linear EST 04-OCT-2001
BI838213
LOCUS 603083131.F1 NIH_MGC_120 Homo sapiens cDNA clone IMAGE:5222574 5',
DEFINITION mRNA sequence.
ACCESSION BI838213 GI:15949763
VERSION BI838213
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 772)
AUTHORS NIH-MGC <http://mhc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.jnl.gov>
Plate: LHAM1559 row: m column: 07
High quality sequence stop: 767.
Location/Qualifiers

FEATURES

source

```

1..772
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5222574"
/lab_host="DH10B"
/clone_1lb="NIH_MGC_120"
/note="Organ: pooled pancreas and spleen; Vector:
pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of spleen and pancreas from 28 yo
male. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 1.5 kb, insert size range 1-2.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 025. Note: this is a NIH_MGC Library."

```

ORIGIN

Alignment Scores:

Pred. No.:	2,996-51	Length:	772
Score:	620.00	Matches:	126
Percent Similarity:	79.64%	Conservative:	7
Best Local Similarity:	75.45%	Mismatches:	21
Query Match:	25.40%	Indels:	13
DB:	4	Gaps:	2

US-09-596-774-6 (1-461) x BI838213 (1-772)

```

QY 296 HisProthrglyThrsSerGlnProGlnArgProGlnuSpCySargProArgGlySerVal 315
DB 90 CATCTCGAGGACAGATTGCCGATTACAGA-----119
QY 316 LysGlyThrGlyLeu-AspPheLeuGluAspProLysLeuGlySerThrLeuAspGlyI 335
DB 120 ---GGCACAGAGCTTTGGCTGCTG---GATCCCAAACTGCTACCTGCTGAGTGAAT 173
QY 335 eLeuPheIleTyrglyValIleIleThrAlaLeuTyrrLeuArgAlaLysPheSerArgse 355
DB 174 CTTCTTCATCATAGTGTGATCTTCTCAGTCTTGTCTGAGTGAAGTTCACAGAGAG 233
QY 415 sMetAlaGlnAlaTyrrSerGlnIleGlyThrLysGlyGluArgArgGlyLysGlyH 435
DB 354 GATGGCGAGGCTTACAGTGAATTTGGGATGAAAGCGCGGAGGGGCAAGGGGCA 413
QY 435 sAspGlyLeuTyrrGlnGlyLeuSerThrAlaThrLysAspThrTyrrAspAlaLeuH 455
DB 474 CGATGGCTTTTACAGGCTCTCAGTACAGCACAAGGACCTTACGAGCCCTTCACAT 533

```

OY 455 tGlnThrLeuAlaProArg 461
 DB 534 GCAGGCCCTGCCCTCGC 552

RESULT 13
 BM921412 1038 bp mRNA linear EST 12-MAR-2002
 LOCUS AGENCOURT 6626432 NIH_MGC_115 Homo sapiens CDNA clone IMAGE:5752844
 DEFINITION 5', mRNA sequence.
 BM921412
 VERSION BM921412.1 GI:19371791
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 1 (bases 1 to 1038)
 NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cga@bbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNLN at:
 http://image.lnl.gov
 Plate: LLM12787 row: k. column: 21
 High quality sequence stop: 730.

FEATURES
 source location/Qualifiers

1..1038
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5752844"
 /lab_host="MDH108"
 /clone_lib="NIH_MGC_115"
 /note="Organ: pooled brain, lung, testis; Vector: pcwv-SPOrTE; Site 1: NotI; Site 2: EcoRV (destroyed); RNA male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH_MGC Library."

ALIGNMENT SCORES:

Pred. No.: 4.56e-51 Length: 1038
 Score: 620.00 Matches: 126
 Percent Similarity: 79.64% Conservative: 7
 Best Local Similarity: 75.45% Mismatches: 21
 Query Match: 25.40% Indels: 13
 DB: 5 Gaps: 2

US-09-596-774-6 (1-461) x BM921412 (1-1038)

OY 296 HlsProthrglYThSerglnProglInArpProgluaAspCyargProArgrlySerVal 315
 DB 86 CATCTTCAGGACAGATTGGCCGATTACAGA----- 115
 OY 316 LysGlyThrGlyLeu-AspPheLeuGluAspProlyLeuCyArgTyrlLeuLeuAspGly11 335
 DB 116 ---GGCACAGAGCTTTGGCTGCTG---GATCCCAAATCTGCTACCTGCTGATGAAT 169
 OY 335 eleuPhelellyrGlyVal11lelThrAlaLeuTyrlLeuAlaIalysPheSerArgSe 355
 DB 170 CCTCTCATCATATGATGATTCACATTCACGCTTGTCTGAGAGTGAAGTTACACAGAG 229

OY 355 rAlaGlnThrAlaAlaasnLeuGlnAspProAsnGlnLeuTyrlXhnglnLeuAsnLeuG1 375
 DB 230 CGCAGAGCCCCCGGTACGACAGAGGCCAGAACCGCTCATATACGAGCTCAATCTAG 289
 OY 375 YArGArgGlnGlnTyrlAspVal1leuGlnLysLysAlaArgAlaArgAspProGlnumetGly1 395
 DB 290 ACGAAGAGAGAGATACGATGTTTGGACMACAGACGTGGCCGGGACCTTGAGATGGGGG 349
 OY 395 YLysGlnGlnArgArgArgAspProGlnGlnGlyVal1TyrlAsnAlaLeuGlnLysAspLy 415
 DB 350 AAAGCCCGACAGAGAGAAACCTTCAGAGAGCTGTACATGACTGCAGAAATATA 409
 OY 415 sMetAlaGlnAlaTyrlSerglnLeuGlyThrLysGlyGlnArgArgArgGlyLys1 435
 DB 410 GATGGCGGAGGCTTACAGTGAATTTGGGATGAAGGGAGGAGCGCGGAGGCAAGGCA 469
 OY 435 sAspGlyLeuTyrlGlnGlyLeuSerrThrAlaThLysAspThTyrlAspAlaLeuH1me 455
 DB 470 CGATGGCCTTTACACAGGTTCTCACTACAGCCACCAAGGACCTTACGAGCCCTTCAAT 529

OY 455 tGlnThrLeuAlaProArg 461
 DB 530 GCAGGCCCTGCCCTCGC 548

RESULT 14
 AL554350 1113 bp mRNA linear EST 30-MAR-2004
 LOCUS AL554350 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
 DEFINITION clone CSOD1082Y123 5-PRIME, mRNA sequence.
 AL554350
 VERSION AL554350.3 GI:45859108
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
 1 (bases 1 to 1113)
 Li, W.B., Gruber, C., Jesssee, J. and Polayres, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished (2001)
 COMMENT On Feb 15, 2001 this sequence version replaced gi:31276162.
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pcwvSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 484.r

For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna/ys=CSOD1082AE120P1&c=484.r.

FEATURES

source location/Qualifiers

1..1113
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CSOD1082Y123"
 /tissue type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pcwvSPORT 6 vector. Library was normalized."

ORIGIN

Alignment Scores:
 Pred. No.: 5.04e-51 Length: 1113
 Score: 620.00 Matches: 126
 Percent Similarity: 79.64% Conservative: 7
 Best Local Similarity: 75.45% Mismatches: 21
 Query Match: 25.40% Indels: 13

DB: 1 Gaps: 2

US-09-596-774-6 (1-461) x AL554350 (1-1113)

QY 296 HisProThrGlyThrSerGlnProGlnArgProGlnuAspCysArgProArgGlySerVal 315
 Db 87 CATCTCGAGCGACAGATTGCCGATTACAGA-----116

QY 316 LysGlyThrGlyLeu-AspPheLeuGluAspProLysLeuGlySerVal 335
 Db 117 ---GGCAGAGAGCTTTGGCTCTG---GATCCAAATCTGCTACCTGCTGATGGAT 170

QY 335 eleuPheIleuTyrglyValIleIleThrAlaLeuTyrlleuArgAlaLysPheSerArgSe 355
 Db 171 CCTCTTATCTATGCTGCTCATCTTCTACCTGCTTGTCTTGAGAGTTCAGACAGAG 230

QY 355 fAlaGluThrAlaAlaAsnLeuGlnAspProAsnGlnLeuTyrlleuAsnLeuG1 375
 Db 231 CGAGAGCGCCCGCGTACAGAGGCGCAGAACCACTATACAGAGCTCAATCTAGG 290

QY 375 YArgArgGluGluTyrlleuValleuGluLysLysArgAlaArgAspProGluMetGlyG1 395
 Db 291 ACGAAGAGAGAGTACGATGTTTGGACAGAGAGCGGCGGAGCCCTGATGGGGGG 350

QY 395 YLysGlnGlnArgArgAspProGlnGluGlyValTyrlleuAlaLeuGlnLysAspLy 415
 Db 351 AAAGCCGATTA 410

QY 415 sMetAlaGluAlaTyrlleuSerGluIleGlyThrLysGlyGluArgArgGlyLysGlyH1 435
 Db 411 GATGGCGAGAGCGCTACAGTGAATGGGATGAGAGAGAGAGAGAGAGAGAGAGAGCA 470

QY 435 sAspGlyLeuTyrglnGlyLeuSerThrAlaThrLysAspThrTyrlleuAlaLeuHisMe 455
 Db 471 CGATGGCTTTTACAGAGGCTCTCAGTACAGCAGAGAGAGAGAGAGAGAGAGAGAGAT 530

QY 455 tGlnThrLeuAlaProArg 461
 Db 531 GCAGGCGCTGCCCTCTGCC 549

RESULT 15
 CR625818 1546 bp mRNA linear HTC 21-JUL-2004
 LOCUS full-length cDNA clone CS0Dj005YJ16 of T cells (Jurkat cell line)
 DEFINITION Cot-10-normalized of Homo sapiens (human).
 ACCESSION CR625818 GI:50506625
 VERSION CR625818.1
 KEYWORDS HTC; CNSLT cDNA.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 1546)
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 TITLE Full-length cDNA libraries and normalization
 JOURNAL Unpublished
 REMARK Contact: Feng Liang Email: fliang@life.technet.com URL: <http://fulllength.invitrogen.com/InvitrogenCorporation1600>
 Faraday Avenue
 2 (bases 1 to 1546)
 REFERENCE Genoscope.
 AUTHORS Direct Submission
 TITLE Submitted (20-JUN-2004) Genoscope - Centre National de Sequencage :
 JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail: secref@genoscope.cns.fr)
 COMMENT - Web: www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dt) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen.
 Location/Qualifiers
 1. 1546
 /organism="Homo sapiens"

ORIGIN

Alignment Scores:

Pred. No.	Score	Percent Similarity	Best Local Similarity	Query Match	DB
8.06e-51	620.00	79.64%	75.45%	25.40%	3
Length: 1546	Matches: 126	Conservative: 7	Mismatches: 21	Indels: 13	2

US-09-596-774-6 (1-461) x CR625818 (1-1546)

QY 296 HisProThrGlyThrSerGlnProGlnArgProGlnuAspCysArgProArgGlySerVal 315
 Db 69 CATCTCGAGCGACAGATTGCCGATTACAGA-----98

QY 316 LysGlyThrGlyLeu-AspPheLeuGluAspProLysLeuGlySerVal 335
 Db 99 ---GGCAGAGAGCTTTGGCTCTG---GATCCAAATCTGCTACCTGCTGATGGAT 152

QY 335 eleuPheIleuTyrglyValIleIleThrAlaLeuTyrlleuArgAlaLysPheSerArgSe 355
 Db 153 CCTCTTATCTATGCTGCTCATCTTCTACCTGCTTGTCTTGAGAGTTCAGACAGAG 212

QY 355 fAlaGluThrAlaAlaAsnLeuGlnAspProAsnGlnLeuTyrlleuAsnLeuG1 375
 Db 213 CGAGAGCGCCCGCGTACAGAGGCGCAGAACCACTATACAGAGCTCAATCTAGG 272

QY 375 YArgArgGluGluTyrlleuValleuGluLysLysArgAlaArgAspProGluMetGlyG1 395
 Db 273 ACGAAGAGAGAGTACGATGTTTGGACAGAGAGAGAGAGAGAGAGAGAGAGAT 332

QY 395 YLysGlnGlnArgArgAspProGlnGluGlyValTyrlleuAlaLeuGlnLysAspLy 415
 Db 333 AAAGCCGATTA 392

QY 415 sMetAlaGluAlaTyrlleuSerGluIleGlyThrLysGlyGluArgArgGlyLysGlyH1 435
 Db 393 GATGGCGAGAGCGCTACAGTGAATGGGATGAGAGAGAGAGAGAGAGAGAGAGAGCA 452

QY 435 sAspGlyLeuTyrglnGlyLeuSerThrAlaThrLysAspThrTyrlleuAlaLeuHisMe 455
 Db 453 CGATGGCTTTTACAGAGGCTCTCAGTACAGCAGAGAGAGAGAGAGAGAGAGAGAGAT 512

QY 455 tGlnThrLeuAlaProArg 461
 Db 513 GCAGGCGCTGCCCTCTGCC 531

Search completed: April 20, 2005, 16:33:39
 Job time : 4534.31 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2005, 06:50:16 ; Search time 65.7843 Seconds
(without alignments)
3588.524 Million cell updates/sec

Title: US-09-596-774-6
Perfect score: 2441
Sequence: 1 MAWWTLLFLPMAAKVPKQI.....LSTATKDTYDALHMQTLAPR 461

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Uniprot_03.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	904	37.0	241	Q921A6	Q921A6 mus musculu
2	832	34.1	244	Q65ZC8	Q65ZC8 homo sapien
3	778	31.9	240	Q65ZC9	Q65ZC9 homo sapien
4	776	31.8	255	Q6KBD5	Q6KBD5 mus musculu
5	767	31.4	243	Q7TQM2	Q7TQM2 mus musculu
6	760.5	31.2	487	Q65ZL2	Q65ZL2 mus sp. fv/
7	748	30.6	298	Q9QYF0	Q9QYF0 synthetic c
8	746.5	30.6	218	Q925S1	Q925S1 mus musculu
9	725.5	29.7	248	Q65Z07	Q65Z07 mus sp. b3(
10	713	29.2	164	CD3Z_MOUSE	CD3Z_MOUSE
11	618	25.3	164	CD3Z_HUMAN	P20963 homo sapien
12	605	24.8	206	CD3H_MOUSE	P20920 mus musculu
13	602.5	24.7	163	CD3Z_PIG	Q9X819 sus scrofa
14	589.5	24.1	165	CD3Z_RABIT	Q9X819 sus scrofa
15	569	23.3	484	Q99LAK6	Q99LAK6 mus musculu
16	539	22.1	166	CD3Z_SHEEP	P23329 ovis aries
17	498	20.4	140	Q65ZL3	Q65ZL3 mus sp. t91
18	486	19.9	149	KVSA_MOUSE	P01633 mus musculu
19	454.5	18.6	116	Q683Y7	Q683Y7 mus musculu
20	438.5	18.0	125	Q6PIL0	Q6PIL0 homo sapien
21	431.5	17.7	497	Q6WY24	Q6WY24 homo sapien
22	430	17.6	518	Q6N030	Q6N030 homo sapien
23	429.5	17.3	480	Q6P089	Q6P089 homo sapien
24	422.5	17.3	480	Q6P089	Q6P089 homo sapien
25	419	17.2	102	Q9JL79	Q9JL79 mus musculu
26	417.5	17.1	500	Q6N091	Q6N091 homo sapien
27	417	17.1	108	Q8VJ10	Q8VJ10 mus musculu
28	414	17.0	500	Q9BRV0	Q9BRV0 homo sapien
29	413	16.9	99	Q9JL74	Q9JL74 mus musculu
30	410	16.8	498	Q6N041	Q6N041 homo sapien
31	409	16.8	170	Q925S2	Q925S2 mus musculu

32	404.5	16.6	469	2	Q7Z7P5	Q7Z7P5 homo sapien
33	396	16.2	470	2	Q7TMM1	Q7TMM1 mus musculu
34	395	16.2	473	2	Q9D8L4	Q9D8L4 mus musculu
35	391.5	16.0	236	2	Q6GMX8	Q6GMX8 homo sapien
36	390	16.0	136	1	KVSB_MOUSE	P01634 mus musculu
37	390	16.0	159	2	Q96Q50	Q96Q50 homo sapien
38	389.5	16.0	481	2	Q91WT1	Q91WT1 mus musculu
39	389	15.9	108	1	KVLY_HUMAN	P80362 homo sapien
40	388	15.9	480	2	Q8K0Z4	Q8K0Z4 mus musculu
41	387.5	15.9	168	2	Q8VDC9	Q8VDC9 mus musculu
42	387.5	15.9	236	2	Q6PH17	Q6PH17 homo sapien
43	387.5	15.9	322	2	Q6KAV0	Q6KAV0 homo sapien
44	386.5	15.8	236	2	Q6GMW1	Q6GMW1 homo sapien
45	386	15.8	613	2	Q8VCX7	Q8VCX7 mus musculu

ALIGNMENTS

RESULT 1

ID	Q921A6	PRELIMINARY:	PRT:	241 AA.
AC	Q921A6			
DT	01-DEC-2001 (TREMBLrel. 19, Created)			
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)			
DE	Anti-CEA 79 single chain Fv (fragment).			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=98170165; PubMed=9509426;			
RA	Chung J.H., Choi S.-J., Kim H.-J., Kim I.-J., Choi I.H., Lee S.-D.,			
RA	Yi K.S., Suh P.-G., Ryu S.H., Chung H.K.;			
RT	"Cloning and characterization of cDNAs encoding VH and VL of a			
RT	monoclonal anti-CEA antibody (CEA 79) cross-reactive with NCA-95 and			
RT	generation of a single-chain Fv molecule (scfv).";			
RL	MoJ. Cells 7:816-818(1997).			
DR	EMBL; U88067; AAB48044.1; -.			
DR	PIR; S19965; S19965.			
DR	PIR; S19967; S19967.			
DR	PIR; S19968; S19968.			
DR	HSSP; S26325; S26325.			
DR	HSSP; P01607; IBMW.			
DR	SMART; SM00406; IGV; 2.			
DR	PROSITE; PS50835; IG_LIKE; 2.			
FT	NON_TER 1			
FT	NON_TER 241			
SO	SEQUENCE 241 AA; 26086 MW; 0276887248B9C771 CRC64;			
Query Match	37.0%; Score 904; DB 2; Length 241;			
Best Local Similarity	70.2%; Pred. No. 1.7e-58;			
Matches	172; Conservative 29; Mismatches 34; Indels 10; Gaps 4;			
Oy	19 QIOLVSGPELKKRGKGTYSKASGYPTNYGNNWQAQACGLKMMGWTNTGSTSTF 78			
Oy	1 QVQLKQSGPELKKRGKGTYSKASGYPTDYGMNWKQAQAGKGLKMMGWTNTYGEPT 60			
Oy	79 ADPKGRFDFSLFETASNTAYLQINNLKSEDMATYFCARMEVYHGVVPWGQGTITVYSSG 138			
Oy	61 ADPKGRFDFSLFETASNTAYLQINNLKSEDMATYFCARMDLR-YFDWGGQGTITVYSSG 119			
Oy	139 GGGSGGGSGGGSGGGSDIQLTOSHKFLSTSVGDRVSIITCKASQDVYNAVAVQAQKQSPK- 197			
Oy	120 GGGSGGGSGGGSGGGSDIQLTOSHKFLSTSVGDRVSIITCKASQDVYNAVAVQAQKQSPK- 179			
Oy	198 ---LTIYSASRRYGVSRFTGSSGSDPTFTTISVQAEDLAVYFCQOHPTPTPTSGGT 254			
Oy	180 AHTHTIYI---QGIPIRSRSGSGSDPTFTTISVQAEDLAVYFCQOHPTPTPTSGGT 234			
Oy	255 KLEIRK 259			

Db 235 KLEIK 239

RESULT 2

Q652C8 PRELIMINARY; PRT; 244 AA.

AC 0652C8: 25-OCT-2004 (Tremblrel. 28, Created)
 DT 25-OCT-2004 (Tremblrel. 28, Last sequence update)
 DT 25-OCT-2004 (Tremblrel. 28, Last annotation update)
 DE Single-chain Fv (Fragment).
 GN Name=scfv;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;
 RX MEDLINE=9219263;
 RA Kontermann R.E., Wing M.G., Winter G.;
 RT "Complement recruitment using bispecific diabodies";
 RL Nat. Biotechnol. 15:629-631(1997).
 DR EMBL; Y13057; CA73500.1; -.
 DR InterPro; IPR003599; Ig_2.
 DR InterPro; IPR007110; Ig_1like.
 DR InterPro; IPR003596; Ig_v.
 DR Pfam; PF00047; Ig_2.
 DR SMART; SM00409; Ig; 2.
 DR SMART; SM00406; Ig; 2.
 DR PROSITE; PS50835; Ig_Like; 2.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 244 AA; 26127 MW; 4B1F17868338F2BF CRC64;

Query Match 34.1%; Score 832; DB 2; Length 244;
 Best Local Similarity 61.9%; Pred. No. 3.4e-53;
 Matches 151; Conservative 44; Mismatches 45; Indels 4; Gaps 2;

QY 19 QIOLVOSGPELKKRGFTVKISCRASGYPTNYGMNVKQAPGQGLKRMGMINTSGESTF 78
 Db 1 QVQLVOSGAEVKRPGDSVKVSCRASGYTFSQHYHWVRQAPGQGLKRMGMINTSGESTF 60
 QY 79 ADFPKGRFDESLTSANTAYLQINNLKSEDMATYFCAR--WEYHGVPVPGQTTVTS 135
 Db 61 AGRFGKRVMTTRDTISAAVMEVSRSLRSDPTAVYYCARBEGSAIYG-NDVWGCTLVTV 119
 QY 136 SSGGGGGGGGGGGGGSDIQLTQSHKFLSTSVGDRVSITCKASODVYNAVAVYQKPGQS 195
 Db 120 SSGGGGGGGGGGGGGSDIQMTQSPSTLSASIGDRVITTCRASGCIYHMLAWYQKPGKA 179
 QY 196 PKLLIYASASRYTGVPRFTSGSGGPDFTTISVQAEIDLAVYFCQGHFTPTFPFGSGTK 255
 Db 180 PKLLIYASASLASGAPRFSGSGSDFTTITISLQPDPAIYVYCOQYSNYPLTFPGSGTK 239
 QY 256 LEIK 259
 Db 240 LEIK 243

RESULT 3

Q652C9 PRELIMINARY; PRT; 240 AA.

AC 0652C9: 25-OCT-2004 (Tremblrel. 28, Created)
 DT 25-OCT-2004 (Tremblrel. 28, Last sequence update)
 DT 25-OCT-2004 (Tremblrel. 28, Last annotation update)
 DE Single-chain Fv (Fragment).
 GN Name=scfv;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NCBI_TaxID=9606;

RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C1q/7;
 RX MEDLINE=97362799; PubMed=9219263;
 RA Kontermann R.E., Wing M.G., Winter G.;
 RT "Complement recruitment using bispecific diabodies";
 RL Nat. Biotechnol. 15:629-631(1997).
 DR EMBL; Y13056; CA73499.1; -.
 DR InterPro; IPR003599; Ig_2.
 DR InterPro; IPR007110; Ig_1like.
 DR InterPro; IPR003596; Ig_v.
 DR Pfam; PF00047; Ig; 2.
 DR SMART; SM00409; Ig; 2.
 DR SMART; SM00406; Ig; 2.
 DR PROSITE; PS50835; Ig_Like; 2.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 240 AA; 25569 MW; FDCFD34645F64B373 CRC64;

Query Match 31.9%; Score 778; DB 2; Length 240;
 Best Local Similarity 60.3%; Pred. No. 3.1e-49;
 Matches 146; Conservative 39; Mismatches 53; Indels 4; Gaps 2;

QY 19 QIOLVOSGPELKKRGFTVKISCRASGYPTNYGMNVKQAPGQGLKRMGMINTSGESTF 78
 Db 1 QVQLVOSGGLVPGGSLRSLSCASGFTTSSYGMHWVRQAPGQGLKRMGMINTSGESTF 60
 QY 79 ADFPKGRFDESLTSANTAYLQINNLKSEDMATYFCAR--WEYHGVPVPGQTTVTS 137
 Db 61 ADFVKGKFTTISRNDSKNTLYLQWNSLRARDTAVYYCARQW--GDSLDPMGKGLTVTS 117
 QY 138 GGGGGGGGGGGGGSDIQLTQSHKFLSTSVGDRVSITCKASODVYNAVAVYQKPGQSPK 197
 Db 118 GGGGGGGGGGGGGSDIQMTQSPSTLSASIGDRVITTCRASGCIYHMLAWYQKPGKAKR 177
 QY 198 LLIIYASASRYTGVPRFTSGSGGPDFTTISVQAEIDLAVYFCQGHFTPTFPFGSGTKLE 257
 Db 178 LLIIYASASLASRAPSFSGSGSDFTTITISLQPDPAIYVYCOQYSNYPLTFPGSGTKLE 237

RESULT 4

Q6KB05 PRELIMINARY; PRT; 255 AA.

AC 06KB05: 05-JUL-2004 (Tremblrel. 27, Created)
 DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
 DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
 DE scFv B8E5 protein (Fragment).
 GN Name=scFv B8E5;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RX MEDLINE=9219263;
 RA Kontermann R.E., Wing M.G., Winter G.;
 RT "Complement recruitment using bispecific diabodies";
 RL Nat. Biotechnol. 15:629-631(1997).
 DR EMBL; A7746180; CAG34081.1; -.
 DR HSP; P01837; IKCR.
 DR InterPro; IPR003599; Ig.
 DR InterPro; IPR007110; Ig_1like.
 DR InterPro; IPR003596; Ig_v.
 DR SMART; SM00409; Ig; 2.
 DR SMART; SM00406; Ig; 2.
 DR PROSITE; PS50835; Ig_Like; 2.
 FT NON_TER 1
 FT NON_TER 1
 SQ SEQUENCE 255 AA; 27445 MW; B68BD38395DF713B CRC64;

Query Match 31.8%; Score 776; DB 2; Length 255;
 Best Local Similarity 59.8%; Pred. No. 4.8e-49;
 Matches 149; Conservative 36; Mismatches 56; Indels 8; Gaps 2;

QY 19 QIOLVOSGPELKKPGETYKISCKASGYPTNYGMMVVOAPQGLKMMGMINSTGSESTF 78
 DB 1 QVQLQSGSGDLVKKPGASLKVSCAASGFTFSSYGMSWVRGTPDKRLRWATITSGSSTY 60
 QY 79 ADDPGRDPFSLSTANTAYLIQINNLSKEDMATYCAR--EYVGYVPYWGCTTVVS 136
 DB 61 PDSVGRFTTISRDNKNTLYLQMSLSKSEDTAMYYCARHINRYDGAFDYWGCTTVVS 120
 QY 137 SGGSGGGGGSGGGSDIQLTQSHKFLSTVGDRTSITCKASQDYNA-----VAMTQQ 190
 DB 121 SGGSGGGGGSGGGSDIYMAQSPSSLSVAGKIMTCKSSQSLNSRNQNYLAWYQQ 180
 QY 191 KPGQSPKLLIYASSRRTYGVPSRFTGSGSDPFTTISVQAEDLAVYFCQGHFPTPTF 250
 DB 181 KPGQSPKLLIYAGASRRESGVPRFTGSGSGTDFTLTISVQAEDLAVYVCCQNDHSYPLTF 240
 QY 251 GSGTLEIK 259
 DB 241 GAGTLEIK 249

RESULT 5

ID Q7QM2 PRELIMINARY; PRT; 243 AA.
 AC Q7QM2;
 DT 01-OCT-2003 (TREMBLrel. 25, Created)
 DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
 DE Scfv 6H8 protein (Fragment).
 GN Name=Scfv 6H8;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 RX [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=Balb/C;
 RA MEDLINE=22853226; PubMed=12860977; DOI=10.1074/jbc.M306877200;
 RA Peter J.C., Feketevari P., Billiard P., Wallukat G., Hoebeke J.;
 RT the beta-2 adrenergic receptor.";
 RT "scfv single chain antibody variable fragment as inverse agonist for
 RL J. Biol. Chem. 278:36740-36747(2003).
 DR EMBL; AJ574851; CAE00495.1; -.
 DR HSBP; P01751; IAGM.
 DR InterPro: IPR007110; IG-like.
 DR InterPro: IPR003596; IG_v.
 DR SMART; SM00406; IGV; 2.
 DR PROSITE; PS50835; IG_LIKE; 2.
 FT NON TER 1
 SQ SEQUENCE 243 AA; 25976 MW; BFFPF64D2DCAF76 CRC64;

Query Match 31.4%; Score 767; DB 2; Length 243;
 Best Local Similarity 60.6%; Pred. No. 2.1e-48;
 Matches 146; Conservative 42; Mismatches 49; Indels 4; Gaps 3;

QY 19 QIOLVOSGPELKKPGETYKISCKASGYPTNYGMMVVOAPQGLKMMGMINSTGSESTF 78
 DB 1 QVQLQSGSGDLVKKPGASLKVSCAASGFTTFTYMMHWKRGQGLKMMGMINSTGSGITNY 60
 QY 79 ADDPGRDPFSLSTANTAYLIQINNLSKEDMATYCARWEVYHGVYVPYWGCTTVVS 138
 DB 61 DEKFNKGLTVDTSSSTAYMHLSSLASDSAVYCARGG--RG-LDVMGAGTTLTVSSG 117
 QY 139 GGGSGGGGGSGGGSDIQLTQSHKFLSTVGDRTSITCKASQDYNAVAVMYOQKPSQKRL 198
 DB 118 GGGSGGGGGSGGGSDIQMTQSSSSPSVSLGDKVITTCASQSDIYNRLAWYQKPPNARL 177
 QY 199 LYSASRYTGVPSRFTGSGSDPFTTISVQAEDLAVYFCQGHFPTPTFGSGTLEIK 258

DB 178 LISATSLSTGVPSRFGSGGSKDYTLSTLQEDVATVYCCQWSTR--TFGGGTLEIK 236
 QY 259 K 259
 DB 237 K 237

RESULT 6

ID Q65ZL2 PRELIMINARY; PRT; 487 AA.
 AC Q65ZL2;
 DT 25-OCT-2004 (TREMBLrel. 28, Created)
 DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)
 DE FV/M4.
 GN Name=M4-ITN<tau>;
 OS Mus sp.
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10095;
 RX [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=96272580; PubMed=8688499;
 RA Qi Y., Xiang J.;
 RA "A genetically engineered single-gene-encoded anti-TAG72 chimeric
 RT antibody secreted from myeloma cells.";
 RL Hum. Antibodies Hybridomas 6:161-166(1995).
 DR EMBL; S82493; AAB37424.2; -.
 DR InterPro: IPR003599; IG.
 DR InterPro: IPR007110; IG-like.
 DR InterPro: IPR003597; IG_c1.
 DR InterPro: IPR003598; IG_c2.
 DR InterPro: IPR003006; IG_MHC.
 DR InterPro: IPR003596; IG_v.
 DR Pfam; PF00647; CI-sec; 2.
 DR SMART; SM00409; IG; 3.
 DR SMART; SM00407; IGV; 4.
 DR SMART; SM00408; IGV; 2.
 DR SMART; SM00406; IGV; 2.
 DR PROSITE; PS50835; IG_LIKE; 4.
 DR PROSITE; PS00290; IG_MHC; UNKNOWN 1.
 SQ SEQUENCE 487 AA; 53578 MW; CTBA69F3055504 CRC64;

Query Match 31.2%; Score 760.5; DB 2; Length 487;
 Best Local Similarity 47.8%; Pred. No. 1.5e-47;
 Matches 150; Conservative 55; Mismatches 78; Indels 31; Gaps 6;

QY 1 MAMVWTLFLMAAK--VKKQIOLVOSGPELKKPGETYKISCKASGYPTNYGMMVVOAP 59
 DB 1 MEMSWVFLFLSVTTGHSQVQLQSGDLVKKPGASLKVSCAASGFTTDTAIIHMAKQK 60
 QY 60 GQGLKMMGMINSTGSESTFADPFKGRDPFSLSTANTAYLIQINNLSKEDMATYCARWEV 119
 DB 61 EQLLEMGYISPGNDIKYNEKFKGKATLTADKSSSTAYMHLSSLASDSAVYFCR--- 117
 QY 120 YHGVVPYWGCTTVVS SGGSGGGGGSGGGSDIQLTQSHKFLSTVGDRTSITCKASQ 179
 DB 118 --SYGHWGCTTVLT--SGGGGGSGGGSGGGSDIQMTQSSPSVSLVSGELVITTCRASE 174
 QY 180 DVNAVAVMYOQKPSQKRLIYSSSRRTYGVPSRFTGSGSDPFTTISVQAEDLAVYFC 239
 DB 175 NITSNLAWYQKQKSPQLVYATNLADGVPSRFGSGSGSTQYSLKINLSQSEDFSGSY 234
 QY 240 CQGHFPTPTFGSGTLEIKALEISNSVMYFSSVAVYLQKYNSTTTKRVLTSPSPV--- 295
 DB 235 CQHFMTPTPTFGSGTLEIK-----PSKSCDKHTGCP---PGAPBELT 275
 QY 296 -HPTGTSOPARPED 308
 DB 276 GGPVFLFPKPKD 289

```

RESULT 7
ID 090YFO PRELIMINARY; PRT; 298 AA.
AC 090YFO;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE CN 8 single chain antibody.
DE Name=CN 8 scFv;
OS synthetic construct;
OC other sequences; artificial sequences.
OX NCBI_TaxID=32630;
RN (1)
RP SEQUENCE FROM N.A.
RA MEDLINE=20183931; PubMed=10706631; DOI=10.1073/pnas.050582197;
RA Shinohara N., Demura T., Fukuda H.;
RT "Isolation of a vascular cell wall-specific monoclonal antibody
RT recognizing a cell polarity by using a phase display subtraction
RT method.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:2585-2590(2000).
DR EMBL; AB036341; BAAB8633.1; -.
DR PIR; A33933; A33933.
DR PIR; S19112; S19112.
DR HSSP; P01820; 1A70.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR SMART; SM00406; IGV; 2.
DR PROSITE; PS50835; IG LIKE; 2.
SQ SEQUENCE 298 AA; 31867 MW; E0F96B8A17004317 CRC64;

Query Match 30.6%; Score 748; DB 2; Length 298;
Best Local Similarity 52.5%; Pred. No. 6.7e-47;
Matches 145; Conservative 46; Mismatches 77; Indels 8; Gaps 2;

OY 7 LFLFM-----AAAKVKQIOLVOSGPELKKRPETVYISCKASGYPTNYGMNWKAPGQGLKMGWINTSGESTP 59
DB 21 LFLFIPLVPPFYAOPAMAQVQLQDSGGGLVKGSLKSCAASGSDPSRRWMSWVQAQ 80
OY 60 GQGLKMGWINTSGESTPADDFKGRFDFSLTSANTAYLQINNKKSEDMATYFCARWEV 119
DB 81 GKGLFWIGELINPDSSTINYPSTLKDKFLISRDNKNTLYLQMSKVRSEDTALYICAR-AS 139
OY 120 YHGVVPWVGQCTTVTVSSGGGGGGGGGGGGSDIQLTOSHKFLSTSVGDRVSTTCASQ 179
DB 140 YHGSAWVGQCTTVTVSSGGGGGGGGGGGGSDIELTOSASISASVGEVITTCRASG 199
OY 180 DVYNAVAVMYOQKPGQSPKLLIYSASRYTGVPSPRFTSGSGPDFTFTISVQAEIDLAVF 239
DB 200 NINHYLAWYQOKQKSPQLLVYNAKTLADGVPSKFSGSGSGSTQYSLKINSLQPEDFGSY 259
OY 240 CQHFRTPTFTFGSGTKLEIKALISNVMYFSSVVP 275
DB 260 CQHFMTPTFTFGSGTKLEIKRAAGAPVPYDPLEP 295

RESULT 8
ID 0925S1 PRELIMINARY; PRT; 218 AA.
AC 0925S1;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE MKP5 (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A.
RA STRAIN=BALB/c;
RA PubMed=11819679;
RA Cui D., Zeng G., Yan X., Wang F., Tian F., Ren D., Zhao T., Li X.,

```

```

RA Su C.;
RT "Mechanism of exogenous nucleic acids and their precursors improving
RT the repair of intestinal epithelium after irradiation in mice.";
RL World J. Gastroenterol. 6:709-717(2000).
RN (2)
RP SEQUENCE FROM N.A.
RA STRAIN=BALB/c;
RA Cui D., Zeng G., Yan X., Li X., Su C.;
RT "Cloning of mouse gene related to repairing of intestinal epithelium
RT of the irradiated mice by treatment with the intestinal RNA of mice of
RT the same strain.";
RL Int. J. Radiat. Biol. Relat. Stud. Phys. Chem. Med. 19:71-80(2001).
DR EMBL; AF240168; AAK43733.1; -.
DR HSSP; P01665; 1ONZ.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG LIKE; 1.
FT NON TER 218
SQ SEQUENCE 218 AA; 23013 MW; 527E4FA8F7982817 CRC64;

Query Match 30.6%; Score 746.5; DB 2; Length 218;
Best Local Similarity 66.2%; Pred. No. 5.8e-47;
Matches 141; Conservative 28; Mismatches 39; Indels 5; Gaps 2;

OY 19 QIOLVOSGPELKKRPETVTKISCKASGYPTNYGMNWKAPGQGLKMGWINTSGESTP 78
DB 3 QVXLQSGPELKKRPETVRISCKASGYTTAGMQVQKMPGGLKMGWINTSGVPKY 62
OY 79 ADFKGRFDFSLTSANTAYLQINNKKSEDMATYFCARWEVHYGYPVYQCTTVTVSSG 138
DB 63 ABEFKGRFAPSLTSTALQISLNKNEDTATVFCMRD-YDGGPVAWQCTTVTVSSG 121
OY 139 GGGSGGGGGGGGGSDIQLTOSHKFLSTSVGDRVSTTCASODVYNA----VAVYQKPGQ 194
DB 122 GGGSGGGGGGGGGSDIQLTOSPASLAVSLGQRTICRASVSNDIGISFNMWFOQKPGQ 181
OY 195 SPKLLIYSASRYTGVPSPRFTSGSGPDFTFT 227
DB 182 PPKLLIYAAASKQSGVPAGLLASGSGTDFSLNI 214

RESULT 9
ID 065Z07 PRELIMINARY; PRT; 248 AA.
AC 065Z07;
DT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE B3 (Fv)-PB40 (Fragment).
OS Name=B3 (Fv)-PB40;
OC Mus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10095;
RN (1)
RP SEQUENCE FROM N.A.
RA MEDLINE=92020904; PubMed=1924323;
RA Brinkmann U., Pal L.H., Fitzgerald D.J., Willingham M., Pastan I.;
RT "B3 (Fv)-PB38KDEL, a single-chain immunotoxin that causes complete
RT regression of a human carcinoma in mice.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:8616-8620(1991).
DR EMBL; SS7990; AAB1971.2; -.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PR00047; Ig; 2.
DR SMART; SM00406; IGV; 2.
DR SMART; SM00406; IGV; 2.
DR PROSITE; PS50835; IG LIKE; 2.
FT NON TER 248
SQ SEQUENCE 248 AA; 26634 MW; 7A3759B43E570950 CRC64;

Query Match 29.7%; Score 725.5; DB 2; Length 248;
Best Local Similarity 55.3%; Pred. No. 2.4e-45;

```

Matches 136; Conservative 42; Mismatches 63; Indels 5; Gaps 1;

QY 20 IQLVSGSELKKPGTIVKISCAGYPTNMGMMVKQAPGGLKMGMINSTSESTFA 79
 Db 3 VKLVESGGGLVOPGGSLKISCATSGFTSDYMYVWRQPERRLVMAVISNDSSAAYS 62
 QY 80 DDFKRGFPFSLSTANTAVYLQINNKSEDMATYFCARNEVHYGVYGGCTTVVSSGG 139
 Db 63 DTVGRFTISRDMNANTLYLQMSRLKSEDTALYSACRGLAMCAWPAVYGGGLTVVSSGG 122
 QY 140 GSGGGGGGGGGSDIQLTQSHKFLSTSVGRVSTICKASQDVYNA-----VAMYQKPGQ 194
 Db 123 GSGGGGGGGGGSDVLTMTQSLPLSPVSLGDQASISCRSSQIIVHNGNTYLEWLYQKGGQ 182
 QY 195 SPKLLIYASRRYTVPSRFTGSSGDPDTFTISSVQABDLAVYFCQGHFRTPFGSGT 254
 Db 183 SPKLLIYKSNRPFSGVDPDRFSGSGGTDTFTLLKISRVEADLGVYVCFQGSHPVFTFGGT 242
 QY 255 KLEIKA 260
 Db 243 KLEIKA 248

RESULT 10
 ID CD32_MOUSE STANDARD; PRT: 164 AA.
 AC P24161, OSRD3G3;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE T-cell surface glycoprotein CD3 zeta chain precursor (T-cell receptor
 T3 zeta chain).
 GN Name=cd3z; Synonyms=Tcrz;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=86145643; PubMed=3378377;
 RA Weisman A.M., Baniyash M., Hou D., Samelson L.E., Burgess W.H.,
 RT Klausner R.D.;
 RL "Molecular cloning of the zeta chain of the T cell antigen receptor.";
 RL Science 239:1016-1021(1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RP TISSUE=Liver;
 RX MEDLINE=89327299; PubMed=2787796;
 RA Baniyash M., Hsu V.W., Seldin M.F., Klausner R.D.;
 RT "The isolation and characterization of the murine T cell antigen
 receptor zeta chain gene.";
 RL J. Biol. Chem. 264:13252-13257(1989).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Thymus;
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nakaio I., Osato N., Saito R., Suzuki H., Yamanaoka I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,
 RA Baidarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schirral L.M., Kanapin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Brad D., Brucic V., Chochia C., Corbani L.E., Cousins S.,
 RA Dalla E., Dragan T.A., Fletcher C.F., Forrest A., Frazer K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Gotslik A., Gough J.,
 RA Grimmond S., Guelinich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Matsuda L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numa K., Okido T., Pavan W.J., Petrea G., Pesole G.,
 RA Petrovsky N., Piliat R., Pontius J.U., Qi D., Ramachandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sadelain A., Schneider C., Sempke C.A., Secou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavalon M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazune N., Sato K.,
 RA Shireki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shingawa A.,
 RA Yaunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.,
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs";
 RL Nature 420:563-573(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Hematopoietic;
 RX MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Dietchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Umed T.B., Toshilyski S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Adamson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalón D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Paray J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Binkley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [5]
 RP PARTIAL SEQUENCE.
 RX MEDLINE=90239005; PubMed=2139725;
 RA Jin J.J., Clayton L.K., Howard F.D., Koyasu S., Stieh M.,
 RA Steinbrich R., Tarr G.E., Reinherz E.L.;
 RT "Molecular cloning of the CD3 zeta subunit identifies a CD3 zeta-
 RT related product in thymus-derived cells.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:3319-3323(1990).
 RN [6]
 RP ALTERNATIVE SPLICING.
 RX PubMed=2150596;
 RA Ohno H., Saito T.;
 RT "CD3 zeta and eta chains are produced by alternative splicing from a
 RT common gene.";
 RL Int. Immunol. 2:1117-1119(1990).
 RN [7]
 RP ERRATUM.
 RA Ohno H., Saito T.;
 RL Int. Immunol. 4:1339-1339(1992).
 RN [8]
 RP INTERACTION WITH SLA.
 RX MEDLINE=20130290; PubMed=10662792;
 RA Soehnlewski T., Pandey A., Dixit V.M., Weiss A.;
 RT "Src-like adaptor protein (SLAP) is a negative regulator of T cell
 RT receptor signaling";
 RL J. Exp. Med. 191:463-474(2000).
 RN [9]
 RP INTERACTION WITH SLA2.
 RX MEDLINE=22011997; PubMed=11891219; DOI=10.1074/jbc.M10318200;
 RA Pandey A., Ibarola N., Kratchmarova I., Fernandez M.M.,
 RA Constantinescu S.N., Ohara O., Sawasdikosol S., Lodish H.F., Mann M.;
 RT "A novel Src homology 2 domain-containing molecule, Src-like adapter
 RT protein-2 (SLAP-2), which negatively regulates T cell receptor
 RT signaling";
 RL J. Biol. Chem. 277:19131-19138(2002).
 CC -I- FUNCTION: Probable role in assembly and expression of the TCR
 CC complex as well as signal transduction upon antigen triggering.

CC -1- SUBUNIT: Interacts with DOCK2 (By similarity). The TCR/CD3 complex
 CC of T lymphocytes consists of either a TCR alpha/beta or TCR
 CC gamma/delta heterodimer coexpressed at the cell surface with the
 CC invariant subunits of CD3 labeled gamma, delta, epsilon, zeta, and
 CC eta. CD3-zeta forms either homodimers or heterodimers with CD3-
 CC eta. Interacts with SLA and SLA2.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=CD-3-zeta;
 CC IsoId=P24161-1; Sequence=Displayed;
 CC Name=CD-3-eta;
 CC IsoId=P29020-1; Sequence=External;
 CC -1- PFM: Phosphorylated on Tyr residues after T-cell receptor
 CC triggering.
 CC -1- SIMILARITY: Belongs to the CD3z/FCER1g family.
 CC -1- SIMILARITY: Contains 3 ITAM domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: M19729; AAA0171.1; -;
 DR EMBL: U04967; AAA50301.1; -;
 DR EMBL: AK017904; BAB30997.1; -;
 DR EMBL: BC052824; AAH52824.1; -;
 DR PIR: A40104; A40104.
 DR MGI: MGI:88334; C032.
 DR InterPro: IPR003110; ITAM.
 DR Pfam: PF02189; ITAM; 3.
 DR Alternative splicing: Direct protein sequencing; Phosphorylation;
 DR Receptor; Repeat; Signal; T-cell; Transmembrane.
 FT SIGNAL 1 21
 FT CHAIN 22 164 T-cell surface glycoprotein CD3 zeta
 FT DOMAIN 22 30 Extracellular (Potential).
 FT TRANSMEM 31 51 Potential.
 FT DOMAIN 52 164 Cytoplasmic (Potential).
 FT DOMAIN 69 89 ITAM 1.
 FT DOMAIN 108 129 ITAM 2.
 FT DOMAIN 139 159 ITAM 3.
 FT DISULFID 32 32 Interchain (Potential).
 FT MOD RES 153 153 Phosphotyrosine.
 FT CONFLICT 153 153 Y -> C (in Ref. 3).
 SQ SEQUENCE 164 AA; 18637 MW; 188022035A312831 CRC64;

Query Match 29.2%; Score 713; DB 1; Length 164;
 Best Local Similarity 99.3%; Pred. No. 1.2e-44;
 Matches 138; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 323 LEDPKLCYLDGLIFITGVITITALLYLRKFSRSARETANLQDPNQLYNELNLRREEDYV 382
 DB 26 LLDPRKCYLDGLIFITGVITITALLYLRKFSRSARETANLQDPNQLYNELNLRREEDYV 85
 QY 383 LEKRRADPEWKGKQRRNPQESGVNALOKDKAEYYSITGKGRKKRGHDGLYOGI 442
 DB 86 LEKRRADPEWKGKQRRNPQESGVNALOKDKAEYYSITGKGRKKRGHDGLYOGI 145
 QY 443 STATKDTYDALHMQTLAPR 461
 DB 146 STATKDTYDALHMQTLAPR 164

RESULT 11
 ID CD3z_HUMAN
 AC P20963; Q8TAX4;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE T-cell surface glycoprotein CD3 zeta chain precursor (T-cell receptor
 DE T3 zeta chain).
 GN Name=CD3z; Synonyms=TR3z, TCRz;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OK NCBI_taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=89071765; PubMed=2974162;
 RA Weissman A.M., Hou D., Orloff D.G., Modi W.S., Suarez H.,
 RA O'Brien S.J., Klausner R.D.,
 RT "Molecular cloning and chromosomal localization of the human T-cell
 RT receptor zeta chain: distinction from the molecular CD3 complex";
 RL Proc. Natl. Acad. Sci. U.S.A. 85:9705-9713 (1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Pancreas;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Peingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udutin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loguettano N.A., Peters G.J., Abrazmon R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green S.D., Dickson M.C.,
 RA Rodriguez A.C., Gitlwood J.J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Schein J.E., Jones S.J.M., Skalska U., Smailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [3]
 RP INTERACTION WITH HIV-1 NEF.
 RX PubMed=10224289;
 RA Xu X.-N., Laferriere B., Screaton G.R., Kraft M., Wolf D., Kolanus W.,
 RA Mongkolkeha J., McMichael A.J., Baur A.S.;
 RT "Induction of Fas ligand expression by HIV involves the interaction of
 RT Nef with the T cell receptor zeta chain";
 RL J. Exp. Med. 189:1489-1496 (1999).
 RN [4]
 RP INTERACTION WITH SLA.
 RX MEDLINE=99380595; PubMed=10449770; DOI=10.1073/pnas.96.17.9775;
 RA Tang J., Sawadikowski S., Chang J.-H., Burckoff S.J.;
 RT "SLAP, a dimeric adapter protein, plays a functional role in T cell
 RT receptor signaling";
 RL Proc. Natl. Acad. Sci. U.S.A. 96:9775-9780 (1999).
 RN [5]
 RP INTERACTION WITH DOCK2.
 RX MEDLINE=22165501; PubMed=12176041; DOI=10.1016/S0006-291X(02)00931-2;
 RA Nishihara H., Maeda M., Tsuda M., Makino Y., Sawa H., Nagashima K.,
 RA Tanaka S.;
 RT "DOCK2 mediates T cell receptor-induced activation of Rac2 and IL-2
 RT transcription";
 RL Biochem. Biophys. Res. Commun. 296:716-720 (2002).
 RN [6]
 RP PHOSPHORYLATION SITES TYR-83; TYR-111; TYR-123 AND TYR-142.
 RX PubMed=12522270; DOI=10.1073/pnas.2436191100;
 RA Salomon A.R., Ficarro S.B., Brill L.M., Brinker A., Phung Q.T., Peters E.C.,
 RA Ericson C., Sauer K., Brock A., Horn D.M., Schultz P.G., Peters E.C.;
 RT "Profiling of tyrosine phosphorylation pathways in human cells using
 RT mass spectrometry";
 RL Proc. Natl. Acad. Sci. U.S.A. 100:443-448 (2003).
 RN [7]
 RP STRUCTURE BY NMR OF 136-149.

RX MEDLINE=93201600; PubMed=7680960; DOI=10.1016/0092-8674(93)90405-F;
 RA Wakeman G., Shoelson S.E., Pant N., Cowburn D., Kurlyan J.;
 RT "Binding of a high affinity phosphotyrosyl peptide to the Src SH2
 RT domain: crystal structures of the complexed and peptide-free forms.";
 RL Cell 72:779-790(1993).
 CC -1- FUNCTION: Probable role in assembly and expression of the TCR
 CC complex as well as signal transduction upon antigen triggering.
 CC -1- SUBUNIT: The TCR/CD3 complex of T lymphocytes consists of either a
 CC TCR alpha/beta or TCR gamma/delta heterodimer coexpressed at the
 CC cell surface with the invariant subunits of CD3 labeled gamma,
 CC delta, epsilon, zeta, and eta. CD3-zeta forms either homodimers or
 CC heterodimers with CD3-eta. Interacts with SLA and SLA2. Interacts
 CC with DOCK2. Interacts with HIV-1 Nef protein.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=CD-3-zeta;
 CC IsoId=P20963-1; Sequence=Displayed;
 CC Name=CD-3-eta;
 CC IsoId=P20963-2; Sequence=Not described;
 CC -1- PTM: Phosphorylated on Tyr residues after T-cell receptor
 CC triggering (by similarity).
 CC -1- SIMILARITY: Belongs to the CD3Z/FCER1G family.
 CC -1- SIMILARITY: Contains 3 ITAM domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL collaboration
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; J04132; AAA60394.1; -;
 DR EMBL; BC025703; AAH25703.1; -;
 DR PIR; A31768; A31768.
 DR PDB; 1TCE; NMR; B=136-149.
 DR Genew; HGNC:1677; CD3Z.
 DR H-InvDB; HIX001296; -;
 DR MIM; 186780; -;
 DR GO; GO:0005886; C:plasma membrane; TAS.
 DR GO; GO:0042101; C:T-cell receptor complex; TAS.
 DR GO; GO:0005515; F:protein binding; IPI.
 DR GO; GO:0042803; F:protein homodimerization activity; NAS.
 DR InterPro: IPR003110; ITAM.
 DR Pfam; PF02189; ITAM; 3.
 DR SMART; SM00077; ITAM; 3.
 KM 3D-structure; Alternative splicing; Phosphorylation; Receptor; Repeat;
 KM Signal; T-cell; Transmembrane.
 FT SIGNAL 1 21
 FT CHAIN 22 164 T-cell surface glycoprotein CD3 zeta
 FT chain.
 FT DOMAIN 22 30 Extracellular (Potential).
 FT TRANSMEM 31 51 Potential.
 FT DOMAIN 52 164 Cytoplasmic (Potential).
 FT DOMAIN 69 89 ITAM 1.
 FT DOMAIN 108 129 ITAM 2.
 FT DOMAIN 139 159 ITAM 3.
 FT DISULFD 32 32 Interchain (Potential).
 FT MOD_RES 83 83 Phosphotyrosine.
 FT MOD_RES 111 111 Phosphotyrosine.
 FT MOD_RES 123 123 Phosphotyrosine.
 FT MOD_RES 142 142 Phosphotyrosine.
 FT MOD_RES 153 153 Phosphotyrosine (by similarity).
 FT CONFLICT 60 61 DA -> EP (in Ref. 1).
 FT CONFLICT 101 101 Missing (in Ref. 1).
 SQ SEQUENCE 164 AA; 18696 MW; 9408260374856EE9 CRC64;

Query Match 25.3%; Score 618; DB 1; Length 164;
 Best Local Similarity 84.9%; Pred. No. 1.le-37;
 Matches 118; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

323 LEDPRLCYLIDGILFIYGVIIITALYLAKFSRSADPAAYOGQONQVNLNIGREEYDV 382

DB 26 LLEPRLCYLIDGILFIYGVIIITALFLRFKFSRSADPAAYOGQONQVNLNIGREEYDV 85
 QY 383 LLEKRAPDEMGKQQRARRNOEGVYNALOKDKMAEYSEIGTGERRRGHOSLYGL 442
 DB 86 LDKRGDRPEWGGVPRKRNQESGLVNLQDKMAEYSEIGMGERRRGKHDSLYGL 145
 QY 443 STATKDYDALHMOTLAPR 461
 DB 146 STATKDYDALHMOTLAPR 164
 RESULT 12
 CD3H_MOUSE STANDARD; PRT; 206 AA.
 ID CD3H_MOUSE P23020;
 AC 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE T-cell surface glycoprotein CD3 eta chain precursor (T-cell receptor
 DE T3 eta chain).
 GN Name=CD3z; Synonyms=CD3h;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_Taxid=10090;
 RN [1]
 RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.
 RX MEDLINE=90239005; PubMed=2139725;
 RA Jjin Y.J., Clayton L.K., Howard F.B., Koyasu S., Sieh M.,
 RA Steinblich R., Tarr G.E., Reinherz E.L.;
 RT "Molecular cloning of the CD3 eta subunit identifies a CD3 zeta-
 RT related product in thymus-derived cells";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:3319-3323(1990).
 RN [2]
 RP SEQUENCE OF 144-206 FROM N.A.
 RX MEDLINE=91271358; PubMed=1828694;
 RA Clayton L.K., D'Adamo L., Sieh M., Husey R.E., Koyasu S.,
 RA Reinherz E.L., Howard F.B.;
 RT "CD3 eta and CD3 zeta are alternatively spliced products of a common
 RT genetic locus and are transcriptionally and/or post-transcriptionally
 RT regulated during T-cell development.";
 RL Proc. Natl. Acad. Sci. U.S.A. 88:5202-5206(1991).
 RN [3]
 RP SEQUENCE OF 144-206 FROM N.A.
 RX PubMed=2150596;
 RA Ohno H., Salto T.;
 RT "CD3 zeta and eta chains are produced by alternative splicing from a
 RT common gene";
 RL Int. Immunol. 2:1117-1119(1990).
 RN [4]
 RP ERRATUM.
 RA Ohno H., Salto T.;
 RL Int. Immunol. 4:1339-1339(1992).
 CC -1- FUNCTION: Probable role in assembly and expression of the TCR
 CC complex as well as signal transduction upon antigen triggering.
 CC -1- SUBUNIT: The TCR/CD3 complex of T lymphocytes consists of either a
 CC TCR alpha/beta or TCR gamma/delta heterodimer coexpressed at the
 CC cell surface with the invariant subunits of CD3 labeled gamma,
 CC delta, epsilon, zeta, and eta. CD3-eta can be complexed in a
 CC heterodimeric form with CD3-zeta subunit. CD3-zeta homodimer has
 CC not been observed.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=CD-3-zeta;
 CC IsoId=P29020-1; Sequence=Displayed;
 CC Name=CD-3-eta;
 CC IsoId=P29020-1; Sequence=Not described;
 CC -1- PTM: Phosphorylated on Tyr residues after T-cell receptor
 CC triggering (by similarity).
 CC -1- SIMILARITY: Belongs to the CD3Z/FCER1G family.
 CC -1- SIMILARITY: Contains 3 ITAM domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; M3158; AAA37398.1; -
 DR EMBL; M76711; AAA40403.1; -
 DR PIR; A35900; A35900.
 DR MGD; MGI:98334; CD3Z.
 DR InterPro; IPR003110; ITAM.
 DR Pfam; PF02189; ITAM; 2.
 DR SMART; SM00077; ITAM; 2.
 KW Alternative splicing; Direct protein sequencing; Receptor; Repeat;
 KW Signal; T-cell; Transmembrane.
 FT SIGNAL 1 21
 FT CHAIN 22 206
 FT DOMAIN 22 30 T-cell surface glycoprotein CD3 eta
 FT TRANSMEM 31 51 chain.
 FT DOMAIN 52 206 Extracellular (Potential).
 FT DOMAIN 69 89 Potential.
 FT DOMAIN 108 129 Cytoplasmic (Potential).
 FT DOMAIN 139 159 ITAM 1.
 FT DISULFID 32 32 ITAM 2.
 FT DISULFID 32 32 ITAM 3.
 SQ SEQUENCE 206 AA; 23339 MW; 829256A2CF4E444 CRC64;

Query Match 24.8%; Score 605; DB 1; Length 206;
 Best Local Similarity 99.2%; Pred. No. 1.4e-36;
 Matches 117; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 323 LEDPKCYLLDGLIFLYGVITLALYRAKFSRSASFTANLQDPNOLYNELNGRREYDV 382
 DB 26 LIDPKCYLLDGLIFLYGVITLALYRAKFSRSASFTANLQDPNOLYNELNGRREYDV 85
 QY 383 LEKRRARDPWGGKQKRRNPQEGVYNALQKDMAEAYSEIGTGERRRGKHGDLVQ 440
 DB 86 LEKRRARDPWGGKQKRRNPQEGVYNALQKDMAEAYSEIGTGERRRGKHGDLVQ 143

RESULT 13
 CD3Z_PIG STANDARD; PRT; 163 AA.

AC Q9XST9;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE T-cell surface glycoprotein CD3 zeta chain precursor (T-cell receptor
 DE T3 zeta chain).
 GN Name=CD3Z;
 OS Sus scrofa (pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
 OX NCBI_TaxID=9623;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Minnesota miniature swine;
 RA Jie H.-B., Yim D., Kim Y.B.;
 RT "The molecular cloning of porcine CD3 zeta.";
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Probable role in assembly and expression of the TCR
 CC complex as well as signal transduction upon antigen triggering.
 CC -1- SUBUNIT: The TCR/CD3 complex of T lymphocytes consists of either a
 CC TCR alpha/beta or TCR gamma/delta heterodimer coexpressed at the
 CC cell surface with the invariant subunits of CD3 labeled gamma,
 CC delta, epsilon, zeta, and eta (by similarity).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- PTM: Phosphorylated on Tyr residues after T-cell receptor
 CC triggering (by similarity).
 CC -1- SIMILARITY: Belongs to the CD3Z/FCER1G family.
 CC -1- SIMILARITY: Contains 3 ITAM domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; AF153930; AAD34640.1; -
 DR InterPro; IPR003110; ITAM.
 DR Pfam; PF02189; ITAM; 3.
 DR SMART; SM00077; ITAM; 3.
 KW Phosphorylation; Receptor; Repeat; Signal; T-cell; Transmembrane.
 FT SIGNAL 1 21
 FT CHAIN 22 163
 FT DOMAIN 22 30 T-cell surface glycoprotein CD3 zeta
 FT TRANSMEM 31 51 chain.
 FT DOMAIN 52 163 Extracellular (Potential).
 FT DOMAIN 69 89 Potential.
 FT DOMAIN 107 128 ITAM 1.
 FT DOMAIN 138 158 ITAM 2.
 FT DISULFID 32 32 ITAM 3.
 FT MOD RES 152 152 Phosphotyrosine (By similarity).
 SQ SEQUENCE 163 AA; 18568 MW; 34898620B67167C7 CRC64;

Query Match 24.7%; Score 602.5; DB 1; Length 163;
 Best Local Similarity 84.2%; Pred. No. 1.6e-36;
 Matches 117; Conservative 7; Mismatches 14; Indels 1; Gaps 1;

QY 323 LEDPKCYLLDGLIFLYGVITLALYRAKFSRSASFTANLQDPNOLYNELNGRREYDV 382
 DB 26 LIDPKCYLLDGLIFLYGVITLALYRAKFSRSASFTANLQDPNOLYNELNGRREYDV 85
 QY 383 LEKRRARDPWGGKQKRRNPQEGVYNALQKDMAEAYSEIGTGERRRGKHGDLVQ 442
 DB 86 LIDRRRGRDPWGGKQKRRNPQEGVYNALQKDMAEAYSEIGTGERRRGKHGDLVQ 144
 QY 443 STATKDTYALHMOQLAPR 461
 DB 145 STATKDTYALHMOQLAPR 163

RESULT 14
 CD3Z_RABIT STANDARD; PRT; 165 AA.

AC Q9TUF8;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE T-cell surface glycoprotein CD3 zeta chain precursor (T-cell receptor
 DE T3 zeta chain).
 GN Name=CD3Z;
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OX NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B/J X Chdb:HM;
 RA Isono T., Nishimura M.;
 RT "Rabbit CD3 zeta.";
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Probable role in assembly and expression of the TCR
 CC complex as well as signal transduction upon antigen triggering.
 CC -1- SUBUNIT: The TCR/CD3 complex of T lymphocytes consists of either a
 CC TCR alpha/beta or TCR gamma/delta heterodimer coexpressed at the
 CC cell surface with the invariant subunits of CD3 labeled gamma,
 CC delta, epsilon, zeta, and eta (by similarity).
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- PTM: Phosphorylated on Tyr residues after T-cell receptor
 CC triggering (by similarity).
 CC -1- SIMILARITY: Belongs to the CD3Z/FCER1G family.

-1- SIMILARITY: Contains 3 ITAM domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/or send an email to license@isb-sib.ch>).

CC EMBL: AB035152; BAA6994.1; -.
CC InterPro: IPR003110; ITAM.
CC Pfam: PF02189; ITAM; 3.
CC SMART: SM00077; ITAM; 3.
CC Phosphorylation; Receptor; Repeat; Signal; T-cell; Transmembrane.
CC SIGNAL 1 21 By similarity.
CC CHAIN 22 165 T-cell surface glycoprotein CD3 zeta chain.
FT DOMAIN 22 30 Extracellular (Potential).
FT TRANSMEM 31 51 Potential.
FT DOMAIN 52 165 Cytoplasmic (Potential).
FT DOMAIN 69 89 ITAM 1.
FT DOMAIN 107 128 ITAM 2.
FT DOMAIN 140 160 ITAM 3.
FT DISULFD 32 32 Interchain (Potential).
FT MOD RES 154 154 Phosphotyrosine (By similarity).
SQ SEQUENCE 165 AA; 18773 MW; 3183136130BAA4F5 CRC64;

Query Match 24.1%; Score 589.5; DB 1; Length 165;
Best Local Similarity 81.6%; Pred. No. 1.4e-35;
Matches 115; Conservative 9; Mismatches 14; Indels 3; Gaps 2;

QY 323 LEDPPLCYLILGILFIYGVITITLALYLRAKFSRSAAETANLQDPNQLVNEINLNGREEYDV 382
DB 26 LLDKPLCYLILGILFIYGVITITLALYLRAKFSRGEVPVPSPOCHTOLVNEINLNGREEYDV 85

QY 383 LEKKRARDPEMGKQKQRRRNPOEGVYNALQDKNAEAYSEIGTKE--RRRGKHGDLGYQ 440
DB 86 LDKRRGRDPEMGKQKRRKNPOEGVYNALQDKNAEAYSEIGTKE--RRRGKHGDLGYQ 144

QY 441 GLSTRTKDTYALHMQTLAPR 461
DB 145 GLSATATKDTYDALHMQTLAPR 165

RESULT 15
Q99LA6 PRELIMINARY; PRT; 484 AA.
AC Q99LA6;
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=mix FVB/N; TISSUE=Mammary tumor;
RA MEDLINE=2386257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Straube R.G., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Scheider C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heich F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C., Rana S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullany S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G., RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S., RA Krzyzanski M.I., Skalska U., Smallos D.E., Scherch A., Schein J.E., RA Jones S.J., Maira M.A., "Generation and initial analysis of more than 15,000 full-length human RT and mouse cDNA sequences."
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=mix FVB/N; TISSUE=Mammary tumor;
RA Straube R.G., Feingold E.A., Grouse L.H., Derge J.G., Schuler G.D., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altschul S.F., Zeeberg B., Buetow K.H., Scheider C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heich F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C., Rana S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullany S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
SQ SEQUENCE 484 AA; 52567 MW; 8EA4A9BCF582FA CRC64;

Query Match 23.3%; Score 569; DB 2; Length 484;
Best Local Similarity 40.9%; Pred. No. 1.8e-33;
Matches 157; Conservative 36; Mismatches 79; Indels 112; Gaps 19;

QY 1 MAMWTLTFLMAAK-VPKQIOLVOSGPELKKPPEYKISCKASGYPTNGMMVQAP 59
DB 1 MAMWTLTFLMAAQSIOAQIOLVOSGPELKKPPEYKISCKASGYPTNGMMVQAP 60

QY 60 GQGLKMGWINTSYGESTFADDFKGRFDSLETANTAYIQLNNLKSEDMATYCAR-- 116
DB 61 GKGLKMGWNIETGESVYADDFKGRFASLETASTIHQLNNLKEDTITFCASDY 120

QY 117 -WEVYHGVYVGGGTTVYVSSGGGSGGGGGGGGSDIQLTQSHKPLSTVGDVSTTC 175
DB 121 DYDIY--AMDYWGQGTSTVSS-----ESANPTIYPLTLPALSS--DPVILGC 166

QY 176 KASQDVNNAVAYQKQKQSPKLLIYASASRYTVBPRTFGSGSGPPTFTISSVQEDL 235
DB 167 -----LH--DYFPGTMVWTKSGKXKIT----- 189

QY 236 AVYFCQGHFRTPTFGSG-----TKLEIKALB-----ISNSVYFSSVYPLQKVNST 283
DB 190 TVNF-----PPLASGGKRYMSQTLTPAVECPBESVYKCSVQHDSPV--QELDVN 239

QY 284 TYKEVLRTPSVPHFTGTSOP---QRP-ED-----C-----RPRGSV---KG 317
DB 240 CSGP---TPPEPRTIPSCQSLSLQRPALBDLLIGSDASTITCLNGLRNPAGAVFTWEP 296

QY 318 TGLDPLEDPKL-----CYLLDGIL 336
DB 297 TGDVAVQKAVQNSCGCYSSSVL 320

Search completed: April 20, 2005, 07:07:08
Job time : 68.7843 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2005, 06:48:35 ; Search time 65.1759 Seconds
(without alignments)
2628.807 Million cell updates/sec

Title: US-09-596-774-7

Perfect score: 2345
Sequence: 1 QIQVQSGPELKKRGETVKI.....LSTATKDYDALHMQTLAPR 443

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2345	100.0	461	2	AAR85508 Leader-sc
2	1636.5	69.8	483	2	AAW82315 Chimeric
3	1501.5	64.0	428	2	AAW24027 Single ch
4	1501.5	64.0	443	2	AAW24025 Single ch
5	1308	55.8	615	2	AAW95056 GAL4-DT-8
6	1308	55.8	617	2	AAW95057 GAL4-DT-8
7	1303	55.6	711	2	AAW26980 Fv(FRP5) -
8	1302	55.5	530	2	AAW95053 scFv(FRP5)
9	1302	55.5	651	2	AAW05136 scFv(FRP5)
10	1302	55.5	699	2	AAW05138 scFv(FRP5)
11	1302	55.5	892	2	AAW05140 scFv2(225)
12	1302	55.5	892	2	AAW05143 scFv2(FRP)
13	1302	55.5	892	2	AAW05139 scFv2(FRP)
14	1302	55.5	895	2	AAW05142 scFv2(FRP)
15	1302	55.5	899	2	AAW05144 scFv2(FRP)
16	1302	55.5	1020	2	AAW05141 scFv2(FRP)
17	1299	55.4	637	2	AAW26982 (FRP5) -BT
18	1284	54.8	240	2	AAW85494 Erib-2-sp
19	1284	54.8	240	2	AAW05134 Single ch
20	1275	54.4	241	2	AAW26979 Fv heavy
21	1249	53.3	631	7	ADH3466 CE7-speci
22	1015.5	43.3	443	6	ABG7648 Humanised
23	1015.5	43.3	443	6	ABG74240 Chimeric
24	1013	43.2	242	4	AAW20433 Anti-Fix/
25	1011.5	43.1	473	2	AAW26646 Chimeric

26	1011.5	43.1	514	2	AAW26647 Chimeric
27	981.5	41.9	250	2	AAW02278 741F8 ant
28	981.5	41.9	250	2	AAW29261 Anti-c-er
29	981.5	41.9	250	2	AAW22400 Single ch
30	981.5	41.9	250	2	AAW47012 Single ch
31	981.5	41.9	250	2	AAW53168 741F8 ant
32	981.5	41.9	250	2	AAW80422 Anti-c-er
33	981.5	41.9	250	7	ABW00714 741F8 sFv
34	975.5	41.6	259	7	ADG32322 Mouse scF
35	975.5	41.6	329	7	ADG32359 Precursor
36	963.5	41.1	667	2	AAW39573 Sequence
37	962.5	41.0	651	2	AAW26649 Chimeric
38	962.5	41.0	692	2	AAW26650 Chimeric
39	960.5	41.0	259	7	ADG32324 Mouse scF
40	960.5	41.0	329	7	ADG32361 Precursor
41	930	39.7	444	2	AAW36845 Single ch
42	921	39.3	240	8	ADN16154 Mouse C7
43	917.5	39.1	505	5	AAU72875 Human NKG
44	917	39.1	251	7	ADG32321 Mouse scF
45	917	39.1	320	7	ADG32358 Precursor

ALIGNMENTS

RESULT 1					
AAW85508	16-MAR-1996 (first entry)				
ID	AAR85508 standard; protein; 461 AA.				
XX	AAR85508;				
AC					
DT					
XX					
DE	Leader-scFv(FRP5):lyt-2 hinge;zeta.				
XX					
KW	Single chain antibody; scFv; antibody engineering; antitumour;				
KW	tumour antigen binding; T-cell receptor; cytotoxic T-lymphocyte;				
KW	monoclonal antibody; erbB-2; cancer; cell targeting;				
KW	adoptive immunotherapy.				
OS	Synthetic.				
XX					
XX					
XX					
FT	Key				
FT	Peptide				
FT	Domain				
FT	Peptide				
FT	Domain				
FT	WO9530014-A1.				
XX	09-NOV-1995.				
XX	20-APR-1995; 95WO-EP001494.				
XX	02-MAY-1994; 94EP-00810244.				
XX	(CIBA) CIBA GEIGY AG.				
XX	Groner B, Moritz D;				
XX	WPI; 1995-393085/50.				
DR	N-PSDB; AAT05783.				
XX					
PT	New bifunctional proteins for use in killing tumour cells - contg. a				
PT	tumour antigen binding domain, a hinge region and a zeta chain derived				
PT	from a T-cell antigen receptor.				
XX					
PS	Example 1; Page 29-31; 46pp; English.				
XX					

CC A bifunctional protein (AAR85505) consists of a single chain antibody, scFv(RRP)5, directed against the tumour erbB-2 antigen, a hinge region, and a functional zeta chain obtd. from a T-cell receptor. The protein is expressed in host cells, esp. cytotoxic T-lymphocytes, providing them with a defined tumour cell specificity enabling targeting to defined tumour cells and MHC-unrestricted and MHC-independent tumour destruction in vitro or in vivo

XX Sequence 461 AA;

Query Match 100.0%; Score 2345; DB 2; Length 461;
Best Local Similarity 100.0%; Pred. No. 1,2e-155;
Matches 443; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 QIOLVSGPELKKRGETVKISKASGYPTNYGMWVQAQGGGLKMMGMINSTGSETF 60
DB 19 QIOLVSGPELKKRGETVKISKASGYPTNYGMWVQAQGGGLKMMGMINSTGSETF 78
QY 61 ADDRKGPRDFSLSTANTAYLQINNLSKEDMATYFCARMEYVHGVPYWGQTTVTYSSG 120
DB 79 ADDRKGPRDFSLSTANTAYLQINNLSKEDMATYFCARMEYVHGVPYWGQTTVTYSSG 138
QY 121 GGGSGGGSGGGSDIQLTQSHKFLSTSVGDRVSITCKASQDVYNAVAMYOQKPGQSPKL 180
DB 139 GGGSGGGSGGGSDIQLTQSHKFLSTSVGDRVSITCKASQDVYNAVAMYOQKPGQSPKL 198
QY 181 LIYASASRYTGVPSRFTSGSGSPDFTTSSVOAEDLAVYCCQHFRTPTFGSGTKLEI 240
DB 199 LIYASASRYTGVPSRFTSGSGSPDFTTSSVOAEDLAVYCCQHFRTPTFGSGTKLEI 258
QY 241 KALEISNSVMYFSSVVPVLOKNSVTTTKPVLRTPSPVHPPTGTSOPQREDCRPGSVKGT 300
DB 259 KALEISNSVMYFSSVVPVLOKNSVTTTKPVLRTPSPVHPPTGTSOPQREDCRPGSVKGT 318
QY 301 GLDFLEDPKLCYLLDGLIFIVGVIIITALLYLRAKFSRSAAETAANTQDPNQLYNEINLGRRE 360
DB 319 GLDFLEDPKLCYLLDGLIFIVGVIIITALLYLRAKFSRSAAETAANTQDPNQLYNEINLGRRE 378
QY 361 EYDVLKKRADPEMGKQQRNRNPOEGVYNALQDKMAEAYSEIGTGERRRKGHDGL 420
DB 379 EYDVLKKRADPEMGKQQRNRNPOEGVYNALQDKMAEAYSEIGTGERRRKGHDGL 438
QY 421 YQGLSTATKDTYDALHMOTLAPR 443
DB 439 YQGLSTATKDTYDALHMOTLAPR 461

```

RESULT 2

AAW82315 ID AAW82315 standard; protein; 483 AA.

XX AAW82315;

XX 26-FEB-1999 (first entry)

XX Chimeric CD19/CTCR protein.

XX Single chain; antibody; B-cell marker; CD19; ScFvCD19; gene therapy;
XX retroviral vector; tumour-associated antigen; cancer; immunoglobulin;
XX CD8 alpha chain; hinge region; beta chain; transmembrane domain;
XX intracellular domain; T-cell receptor-CD3 complex.

OS Homo sapiens.

OS Synthetic.

XX DE19720152-A1.

XX 05-NOV-1998.

XX 02-MAY-1997; 97DE-01020152.

XX 02-MAY-1997; 97DE-01020152.

XX

PA (DELB-) DELBRUECK CENT MOLEKULARE MEDIZIN MAX.

XX Schroeder A, Schwenkenbecher J, Bargou R;

XX WPI; 1998-584686/50.

XX N-PSDB; AAV73324.

PT Retroviral vector system for gene therapy - by expression of therapeutic gene in T cells.

XX Claim 9; Page 6; 10pp; German.

CC This sequence represents a novel chimeric CD19/CTCR protein which is composed of an immunoglobulin leader sequence, an anti-CD19 single chain antibody fragment (ScFvCD19), a hinge region from the CD8 alpha chain and the transmembrane and intracellular domain of the T-cell receptor-CD3 complex zeta chain. This sequence is used in a retroviral vector system for gene transfer and expression in primary human T cells which involves a retroviral vector containing CDNA corresponding to a therapeutic gene, a promoter selected from the SV40 immediate early promoter or the human II-2 or MHC I promoter and a selectable marker, and a packaging cell line (i.e. a modified retroviral packaging line for enhanced transfer into primary human T lymphocytes). This vector system in which the CDNA encodes a single-chain antibody (scFv) specific for a tumour-associated antigen can be used for gene therapy of cancer, especially by isolating T cells from the patient's blood, stimulating the T cells in vitro, introducing the retroviral vector into the T cells and returning the T cells to the patient, preferably by intravenous injection

XX Sequence 483 AA;

Query Match 69.8%; Score 1636.5; DB 2; Length 483;
Best Local Similarity 70.3%; Pred. No. 4,2e-106;
Matches 326; Conservative 38; Mismatches 79; Indels 21; Gaps 5;

```

QY 1 QIOLVSGPELKKRGETVKISKASGYPTNYGMWVQAQGGGLKMMGMINSTGSETF 60
DB 20 EVQLQESGAEIVRPGSSVTKISKASGYAFSSWMNVKQRPQGLIEWIGIWDGDDPTNY 79
QY 61 ADDRKGPRDFSLSTANTAYLQINNLSKEDMATYFCARMEV-----HGVVVPYWGQTTV 115
DB 80 NGKPKGATLTLAESSSTAAMQSSLASSEDSAVYFCARRRTTGVGPYVAMDYWGQSTSC 139
QY 116 TVSSGGGGS--GGGSGGGSDIQLTQSHKFLSTSVGDRVSITCKASQDY-YNA---VAV 169
DB 140 TASAKTTPKLBEBEPEARVDILITQSPASLAVALQGRATISCKASQSDYDSDSLNW 199
QY 170 YQKPGQSPKLLIYASASRYTGVPSRFTSGSGSPDFTTSSVOAEDLAVYFCQGHFRT 229
DB 200 YQIIPGQPKLLIYDASNLVSGIIPRPSGSGSGDTFLNIHPEKVDAAATYHCQOSTEDP 259
QY 230 FTSSGSKTKLEIKA-----LEISNSVMYFSSVVPVLOKNSVTTTKPVLRTPSPVHP 279
DB 260 WTFGGGTKLEILEQKLISEEDLNLEISNSVMYFSSVVPVLOKNSVTTTKPVLRTPSPVHP 319
QY 280 TGTSPQRPEDCRPGSVKGTGLDFLEDPKLCYLLDGLIFIVGVIIITALLYLRAKFSRSAAE 339
DB 320 TGTSPQRPEDCRPGSVKGTGLDFLEDPKLCYLLDGLIFIVGVIIITALLYLRAKFSRSAAE 379
QY 340 TAANLQDPNQLYNEINLGRREYDVLKKRADPEMGKQQRNRNPOEGVYNALQDKMA 399
DB 380 TAANLQDPNQLYNEINLGRREYDVLKKRADPEMGKQQRNRNPOEGVYNALQDKMA 439
QY 400 EAYSEIGTGERRRKGHDGLYQGLSTATKDTYDALHMOTLAPR 443
DB 440 EAYSEIGTGERRRKGHDGLYQGLSTATKDTYDALHMOTLAPR 483

```

RESULT 3

AAW24027 ID AAW24027 standard; protein; 428 AA.

XX AAW24027;

XX 25-MAR-2003 (revised)
DT 04-MAR-1998 (first entry)
XX Single chain antigen hybrid receptor partial sequence.
DE Hybrid receptor; single chain antigen; gene therapy; diagnosis;
KW signal conduction; receptor; control region.
XX
OS Synthetic.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FT Peptide 1..4
FT /note= "partial leader sequence"
FT Protein 5..428
FT /note= "single chain antigen hybrid receptor"
FT Region 5..124
FT /label= VH_B_1-8
FT /note= "variable heavy chain region of B1-8 antibody"
FT Region 125..139
FT /label= (GLY_Ser)3
FT /note= "linker region"
FT Region 140..249
FT /label= VL_B_1-8
FT /note= "variable light chain region of B1-8 antibody"
FT Region 250..294
FT /label= CD_8alpha_hinge
FT Domain 295..315
FT /label= transmembrane_domain
FT Domain 331..428
FT /label= cytoplasmic_domain
XX
PN W09720938-A2.
XX
XX 12-JUN-1997.
PD
XX 03-DEC-1996; 96WO-DE002334.
PF
XX 05-DEC-1995; 95DE-01045351.
PR
XX (UYFR-) UNIV FREIBURG KLINIKUM ALBERT-LUDWIGS.
PA
XX Mettelmann R, Kulmburg P, Rosenthal F;
PI WPI; 1997-319784/29.
DR N-PSDB; AAT77139.
DR
XX
XX Cells with hybrid receptor having extracellular and intracellular regions
PT of different origins - useful in gene therapy and diagnosis of tumours.
PT
XX
XX Example 3; Fig 4; 46pp; German.
XX
XX This partial sequence encodes a novel single chain antigen hybrid
CC receptor (HR). The full length receptor sequence can be found in AAM24025
CC and contains an extracellular domain specific for the hapten 4-hydroxy-5-
CC 1000-3-nitrophenyl acetate (NIP) coupled to the CD8-alpha region and the
CC transmembrane and signal-conducting intracellular parts of the CD3-zeta
CC molecule. Such hybrid receptors comprise a receptor part localised on the
CC outside of the cell and specific to a particular signal molecule and a
CC receptor part originating from another receptor, localised on the inside
CC of the cell and capable of setting off a signal inside the cell. The cell
CC should also contain at least one other gene construct with a control
CC region which can interact with the signal sent out by the hybrid receptor
CC and thereby control expression of a transgene bound to this control
CC region. Such cells are useful in gene therapy or for diagnostic purposes.
CC (Updated on 25-MAR-2003 to correct PR field.)
XX
SQ Sequence 428 AA;
Query Match 64.0%; Score 1501.5; DB 2; Length 428;
Best Local Similarity 66.4%; Pred. No. 1e-96;
Matches 297; Conservative 44; Mismatches 79; Indels 27; Gaps 6;

QY 1 QIOLVOSGPELKKPEETVKISCKASGYPTNYGMNWKQADPGGLKMGWINTSGESTP 60
DB 5 QVQLQDSAEIYVKGASVSKSCASGTFSTYMMHWKQRRGRGLEWIGRIDPNSGRTKY 64
QY 61 ADDFKGRFDPSLETSANTAYLIQINNLSKEDMATYFCARWEVY-HGYVYWGQTTVTYSS 119
DB 65 NEKFESKATLIVDRPSSSTAYWQLSLTSEDSAVVYCARVDYSSYPDWGQTTVTYSS 124
QY 120 GGGSGGGSGGGSGGSDIQLTSHKFLSTSVGDRYSITCKASQDYY--NAAWVQQRKQ 176
DB 125 GGGSGGGSGGGSGGSAVGTQ-ESALTTSPGETVLTLCRSSSTGAVTTSNVAWVQEKPDH 183
QY 177 SPKLIIYSASRRYGVPSRFTSGSGSPDFTTISVQAEADAVVFCQOHFPTPPFGSGT 236
DB 184 LFTGLIGTNNRAAGVPRFSGSLIGDKAALTTCAGTDEDAITFCALMTYNNHWVFGGCT 243
QY 237 KLEIKALEISNSVMYFSSVPELQKVNSTTKPVLRTPSPVHPGTSGOPQRPEDCRPRS 296
DB 244 KLTV--LE-----EFTTTKPVLRTPSPVHPGTSGOPQRPEDCRPRS 283
QY 297 VKGTGLDFLEDPKLCYLLDGLFTYGVITTLTYLPAKRSRAETRAANIQDPNLYNEIYL 356
DB 284 VKGTGLDF--DPKLCYLLDGLFTYGVITTLTYLRRKRSRAETRAANIQDPNLYNEIYL 341
QY 357 GRREYDYLEKKRRARDEPMGSKQORRRNPOEGVYNALQKDAEAYSRIGTKERRRGKG 416
DB 342 GRREYDYLEKKRRARDEPMGSKQORRRNPOEGVYNALQKDAEAYSRIGTKERRRGKG 401
QY 417 HDGLYQGLSTATKDTYDALHMQTAPR 443
DB 402 HDGLYQGLSTATKDTYDALHMQTAPR 428
RESULT 4
ID AAM24025
ID AAM24025 standard; protein; 443 AA.
XX
XX AAM24025;
AC
XX 25-MAR-2003 (revised)
DT 04-MAR-1998 (first entry)
XX
XX Single chain antigen hybrid receptor.
DE
XX Hybrid receptor; single chain antigen; gene therapy; diagnosis;
KW signal conduction; receptor; control region.
KW
XX
OS Synthetic.
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FT Peptide 1..19
FT /label= leader_peptide
FT Protein 20..444
FT /note= "single chain antigen hybrid receptor"
FT Region 20..139
FT /label= VH_B_1-8
FT /note= "variable heavy chain region of B1-8 antibody"
FT Region 140..154
FT /label= (GLY_Ser)3
FT /note= "linker region"
FT Region 155..264
FT /label= VL_B_1-8
FT /note= "variable light chain region of B1-8 antibody"
FT Region 265..309
FT /label= CD_8-alpha_hinge
FT Domain 310..330
FT /label= transmembrane_domain
FT Domain 331..444
FT /label= cytoplasmic_domain
XX
PN W09720938-A2.

us-09-596-774-7.apr20.rag

[illegible]

XX	01-NOV-1994;	94EP-00810627.
PR	(WELS/) WELS W.	
PA	Wels W, Fominaya J;	
PI	WPI; 1996-239505/24.	
XX	N-PSDB; AAT29413.	
DR		
XX		
PT	Nucleic acid transfer system for gene therapy, e.g. against cancer -	
PT	includes toxin translocation domain to target nucleic acid to specific	
PT	cell.	
XX		
PS	Claim 7; Page 87-89; 106pp; English.	
XX		
CC	A multidomain protein (AAR95057) comprises a FLAG epitope, the DNA	
CC	binding domain of yeast Gal4, a portion of diaphorase toxin that acts as	
CC	a translocation domain, and a single chain antibody, scFv, of monoclonal	
CC	antibody FRP5 (raised against human tumour cell HER2 antigen) that acts	
CC	as a ligand domain. It is the product of a fusion gene (AAT29413), and is	
CC	used with an effector nucleic acid that comprises e.g. a gene to be	
CC	delivered to a cell and a cognate structure for the Gal4 binding	
CC	domain. This provides a novel means of nucleic acid transfer, suitable	
CC	for gene therapy. (Updated on 27-AUG-2003 to correct OS field.)	
XX		
XX	Sequence 617 AA;	
	Query Match	55.8%; Score 1308; Db 2; Length 617;
	Best Local Similarity	96.5%; Pred. No. 5,1e-83;
	Matches 245; Conservative 3; Mismatches 6; Indels 0; Gaps 0	
QY	1 QIOLVQSGPELKKPGEIVKISCKASGYPTNYGMNVAQAPQGLKMMGMINTSGESTF	60
DB	364 QVQLQQSGPELKKPGEIVKISCKASGYPTNYGMNVAQAPQGLKMMGMINTSGESTF	423
QY	61 ADDRKGKRPDELFEISANTATYQIINNLSKEDMATYFCAMVEYHGVPPYWGOSTTYSVG	120
DB	424 ADDRKGKRPDELFEISANTATYQIINNLSKEDMATYFCAMVEYHGVPPYWGOSTTYSVG	483
QY	121 GGGSGGGSGGGGGSDIQLTQSHKFLSTVGDVRSITCKASQDVNAVAVMYQOKPGQSPKL	180
DB	484 GGGSGGGSGGGGGSDIQLTQSHKFLSTVGDVRSITCKASQDVNAVAVMYQOKPGQSPKL	543
QY	181 LIYASASRYTVPSRFTSGSGDPFTFTISSVQAEADLAVYFCQGHFRTPFTFGSGTKLEI	240
DB	544 LIYASASRYTVPSRFTSGSGDPFTFTISSVQAEADLAVYFCQGHFRTPFTFGSGTKLEI	603
QY	241 KALEISNSVMYFSS 254	
DB	604 KALELDSERRFSA 617	
RESULT 7		
AC	AAR26980	
XX	AAR26980 standard; protein; 711 AA.	
XX	AAR26980;	
XX		
DT	25-MAR-2003 (revised)	
DT	11-FEB-1993 (first entry)	
XX		
DE	Fv(FRP5)-phoA recombinant antibody.	
XX		
KW	Monoclonal antibody; light chain; heavy chain; tumour; phoA;	
XX	variable region; alkaline phosphatase; c-erbB-2; ss.	
OS	Synthetic.	
XX		
FH	Key	Location/Qualifiers
FT	Peptide	1..21
FT		/label= ompA_signal_peptide
FT	Domain	23..141

FT	Peptide	/label= FRP5_heavy_chain_variable_domain
FT		142. 156
FT	Domain	/label= Linker
FT		157. 264
FT	Protein	/label= FRP5_light_chain_variable_domain
FT		265. 711
PN		/label= phoa
XX	EP502812-A1.	
XX		
PD	09-SEP-1992.	
XX		
PF	27-JAN-1992;	92EP-00810056.
XX		
PR	05-FEB-1991;	91EP-00810079.
XX		
PA	(CIBA) CIBA GEIGY AG.	
XX		
PI	Wels WS, Hynes NE, Harwerth I, Groner B, Hardman N, Zwickl M,	
DR	WPI; 1992-302096/37.	
XX	N-PSDB; AAQ28257.	
PT	Recombinant antibodies directed to growth factor receptor C-ERBB-2 - for	
PT	diagnosing and treating tumours expressing C-ERBB-2 e.g. breast or	
PT	ovarian tumours.	
XX		
PS	Disclosure; Page 34-40; 67pp; English.	
XX		
CC	The sequence herein is the single chain recombinant antibody desiganted	
CC	FV(FRP5)-phoa. The alkaline phosphatase gene (phoa) was used as a marker	
CC	gene so that E. coli transformed with the fusion gene could be	
CC	identified. The fusion gene was expressed in E. coli and the antibody was	
CC	extracted. This recombinant antibody can be used for the qualitative and	
CC	quantitative determination of c-erbB-2. This can be used for monitoring	
CC	or in-vivo localisation of tumours overexpressing c-erbB-2. (Updated on	
CC	25-MAR-2003 to correct PN field.)	
XX		
XX	Sequence 711AA;	

Query Match	55.6%	Score 1303;	DB 2;	Length 711;
Best Local Similarity	94.2%	Pred. No. 1.3e-82;		
Matches 245;	Conservative	2;	Mismatches 1;	Indels 12; Gaps 1

QY	1	QIQIVQSGPELKKPEETVAKISCKASGYPTNYGNMNVKQAPQGLKMMGMINTSGESTF	60
DB	23	QVQLQQSGPELKKPEETVAKISCKASGYPTNYGNMNVKQAPQGLKMMGMINTSGESTF	82
QY	61	ADDFKGRDFSLSETANTAYILQINNLKSDMATYCARREYVHGVPYWGQITTYVSSG	120
DB	83	ADDFKGRDFSLSETANTAYILQINNLKSDMATYCARREYVHGVPYWGQITTYVSSG	142
QY	121	GGGSGGGSGGGGGSDIQLTQSHKFLSTSVGRVSIITCKASQDVNVAVMYQOKPQGSPL	180
DB	143	GGGSGGGSGGGGGSDIQLTQSHKFLSTSVGRVSIITCKASQDVNVAVMYQOKPQGSPL	202
QY	181	LIYSASSRYTVVPSRFETGSGSDPTFTITSSVQAEDLAVYRCQHFRTPTFGSGTKLEI	240
DB	203	LIYSASSRYTVVPSRFETGSGSDPTFTITSSVQAEDLAVYRCQHFRTPTFGSGTKLEI	262
QY	241	KALEISNSVMYFSSVVPVLQ 260	
DB	263	KALE-----PVLE 270	

RESULT 8		
AAAR95053		
ID	AAAR95053	standard; protein; 530 AA.
XX	AAAR95053;	
XX		
DT	16-OCT-2003	(revised)
DT	18-AUG-1996	(first entry)

XX	scFv(FRPS) - DELTA-DGAL4 multidomain protein.
DE	
XX	
KW	Nucleic acid transfer system; gene transfer; gene therapy;
KW	cell targeting; multidomain protein; vector; cancer; exotoxin A; DELTA;
XX	single chain antibody; scFv; Gal4.
XX	
OS	Mus; sp.
OS	Pseudomonas; aeruginosa.
OS	Saccharomyces cerevisiae.
OS	Chimeric.
OS	Synthetic.
XX	
FH	Key
FT	Location/Qualifiers
FT	Peptide
FT	1..8
FT	/label= FLAG_epitope
FT	9..17
FT	Peptide
FT	/label= Spacer
FT	18..257
FT	/label= ScFv(FRPS)
FT	258..260
FT	Peptide
FT	/label= Spacer
FT	261..375
FT	Domain
FT	/label= ERA
FT	/note= "amino acids 252-366 of exotoxin-A"
FT	376
FT	Peptide
FT	/label= Spacer
FT	377..522
FT	Domain
FT	/label= GAL4
FT	/note= "amino acids 2-147 of yeast GAL4"
FT	523..530
FT	Peptide
FT	/label= Spacer
FT	/note= "endoplasmic reticulum retention peptide"
XX	
PN	WO9613599-A1.
XX	
PD	09-MAY-1996.
XX	
PF	31-OCT-1995; 95NO-BP004270.
XX	
PR	01-NOV-1994; 94EP-00810627.
XX	
PA	(WELLS/) WELLS W.
PI	Weis W, Fominaya J;
XX	
DR	WPI, 1996-239505/24.
XX	
DR	N-PDB; AAT29409.
XX	
PT	Nucleic acid transfer system for gene therapy, e.g. against cancer -
PT	includes toxin translocation domain to target nucleic acid to specific
PT	cell.
XX	
BS	Claim 7; Page 59-61; 106pp; English.
XX	
CC	A multidomain protein (AAR95053) has a FLAG epitope, a single chain
CC	antibody, scFv, of monoclonal antibody FRPs (raised against human tumour
CC	cell HER2 antigen) that acts as a ligand domain, a non-cytotoxic portion
CC	of Pseudomonas aeruginosa exotoxin A acting as a translocation domain and
CC	the DNA binding domain of yeast Gal4. It is the product of a fusion gene
CC	(AAT29409) and can be expressed in E. coli (resulting in removal of ompA
CC	signal peptide). It is used with an effector nucleic acid that comprises
CC	e.g. a gene to be delivered to a cell and a cognate structure for the
CC	GAL4 DNA binding domain. This provides a novel means of nucleic acid
CC	transfer, suitable for gene therapy. (updated on 16-Oct-2003 to
CC	standardise OS field)
XX	
XO	Sequence 530 AA;

Query Match	55.5%;	Score 1302;	DB 2;	Length 530;
Best Local Similarity	99.2%;	Pred. No. 1.1e-82;		
Matches 242;	Conservative 1;	Mismatches 1;	Indels 0;	Gaps 0;

QY 1 QIOLVOSGPELKKPGETYKISCKASGYPTNYGNMVKQAPQGGLKMMGMINTSTGESETF 60
 DB 18 QVQLQSGPELKKPGETYKISCKASGYPTNYGNMVKQAPQGGLKMMGMINTSTGESETF 77
 QY 61 ADDPKGRDFELETSTANTAYLIQINNLSKEDMATYFCARMEVYHGVPYWGQGTITVSSG 120
 DB 78 ADDPKGRDFELETSTANTAYLIQINNLSKEDMATYFCARMEVYHGVPYWGQGTITVSSG 137
 QY 121 GGGSGGGSGGGGSDIQLTQSHKFLSTSVGDRVSTITCKASODVYNAVMYQOKPGQSPKL 180
 DB 138 GGGSGGGSGGGGSDIQLTQSHKFLSTSVGDRVSTITCKASODVYNAVMYQOKPGQSPKL 197
 QY 181 LIYSASRYTGVPSRFTSGSGSPDFTTISVQAEDLAVYFCQGHFRPTFTFGSGTKLEI 240
 DB 198 LIYSASRYTGVPSRFTSGSGSPDFTTISVQAEDLAVYFCQGHFRPTFTFGSGTKLEI 257
 QY 241 KALE 244
 DB 258 KALE 261

RESULT 9

AAW05136
 ID AAW05136 standard; protein; 651 AA.

XX AC AAW05136;
 XX DT 29-JAN-1997 (first entry)
 XX DE scFv(FRP5)-ETA fusion protein.
 XX KW Single chain antibody; scFv; monoclonal antibody; Mab; EGF; erbB-2;
 KW epidermal growth factor; receptor; plasmid pSM202-5; cancer; therapy;
 KW antitumour; exotoxin A; ETA.
 XX MM Mus; gp.
 OS Pseudomonas; aeruginosa.
 OS Synthetic.
 OS Chimeric.
 XX FT Key
 FT FH Location/Qualifiers
 FT FT 1..21
 FT Peptide
 FT /label= Sig peptide
 FT /note= "ompa signal peptide"
 FT Peptide
 FT 22..38
 FT /label= Spacer
 FT Protein
 FT 39..278
 FT /label= scFv(FRP5)
 FT Peptide
 FT 279..289
 FT /label= Spacer
 FT Protein
 FT 290..651
 FT /label= ETA
 FT /note= "exotoxin A amino acids 252-613"
 XX EP739984-A1.
 XX PD 30-OCT-1996.
 XX PF 26-APR-1995; 95EP-00106275.
 XX PR 26-APR-1995; 95EP-00106275.
 XX PA (SANT-) SAN TUMORFORSCHUNGS GMBH.
 XX PI Wels W, Schmidt M, Groner B;
 XX DR MPI; 1996-478748/48.
 XX DR N-PSDB; AAT42036.
 XX PT Bivalent fusion proteins that bind epidermal growth factor receptor or
 XX PT analogues - and comprise at least two different cell surface binding
 XX PT domain(s), useful for tumour therapy.

PS Example 7; Page 21-22; 52pp; English.

XX scFv(FRP5)-ETA (AAW05136) comprises the single-chain binding region (see
 CC also AAW05134) of murine monoclonal antibody FRP5, which is specific for
 CC human epidermal growth factor receptor erbB-2, joined to exotoxin A
 CC (ETA). It is encoded by plasmid pSM202-5 (see also AAT42036) obt'd. by
 CC ligating an scFv(FRP5) gene (AAT42034) into plasmid pSM200 contg. the
 CC Pseudomonas aeruginosa PAK ETA gene. The construct can be used to produce
 CC novel bivalent fusion proteins (see also AAW05135-44) in bacterial host
 CC cells, for use as antitumour agents
 XX SQ Sequence 651 AA;

Query Match

Beat Local Similarity 55.5%; Score 1302; DB 2; Length 651;
 Matches 242; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QIOLVOSGPELKKPGETYKISCKASGYPTNYGNMVKQAPQGGLKMMGMINTSTGESETF 60
 DB 39 QVQLQSGPELKKPGETYKISCKASGYPTNYGNMVKQAPQGGLKMMGMINTSTGESETF 98
 QY 61 ADDPKGRDFELETSTANTAYLIQINNLSKEDMATYFCARMEVYHGVPYWGQGTITVSSG 120
 DB 99 ADDPKGRDFELETSTANTAYLIQINNLSKEDMATYFCARMEVYHGVPYWGQGTITVSSG 158
 QY 121 GGGSGGGSGGGGSDIQLTQSHKFLSTSVGDRVSTITCKASODVYNAVMYQOKPGQSPKL 180
 DB 159 GGGSGGGSGGGGSDIQLTQSHKFLSTSVGDRVSTITCKASODVYNAVMYQOKPGQSPKL 218
 QY 181 LIYSASRYTGVPSRFTSGSGSPDFTTISVQAEDLAVYFCQGHFRPTFTFGSGTKLEI 240
 DB 219 LIYSASRYTGVPSRFTSGSGSPDFTTISVQAEDLAVYFCQGHFRPTFTFGSGTKLEI 278
 QY 241 KALE 244
 DB 279 KALE 282

RESULT 10

AAW05138
 ID AAW05138 standard; protein; 699 AA.

XX AC AAW05138;
 XX DT 29-JAN-1997 (first entry)
 XX DE scFv(FRP5)/TGF alpha-ETA.
 XX KW Single chain antibody; scFv; monoclonal antibody; Mab; EGF; erbB-2;
 KW epidermal growth factor; receptor; plasmid pMS238-5-TGF; cancer;
 KW exotoxin A; ETA; transforming growth factor alpha; TGF; antitumour.
 XX MM Homo; sapiens.
 OS Mus; gp.
 OS Pseudomonas; aeruginosa.
 OS Synthetic.
 OS Chimeric.
 XX FT Key
 FT FH Location/Qualifiers
 FT FT 1..21
 FT Peptide
 FT /label= Sig peptide
 FT /note= "ompa signal peptide"
 FT Peptide
 FT 22..38
 FT /label= Spacer
 FT Region
 FT 39..278
 FT /label= scFv(FRP5)
 FT Peptide
 FT 279..289
 FT /label= Spacer
 FT Region
 FT 290..404
 FT /label= ETA
 FT /note= "exotoxin A amino acids 252-366"
 FT 405..410
 FT /label= Spacer

FT	Region	411..460
FT	/label= TGF- α ha	
FT	Peptide	461..465
FT	/label= Spacer	
FT	Region	466..699
FT	/label= ETA	
FT	/note="endotoxin-A amino acids 380-613"	
XX		
XX	EP739984-A1.	
PN		
XX	30-OCT-1996.	
XX		
PF	26-APR-1995; 95EP-00106275.	
XX		
PR	26-APR-1995; 95EP-00106275.	
XX		
PA	(SANT-) SAN TUMORFORSCHUNGS GMBH.	
XX		
PI	Wels W, Schmidt M, Groner B;	
XX		
DR	WPI: 1996-478748/48.	
XX	DR N-PSDB; AAT42038.	
PT		
PT	Bivalent fusion proteins that bind epidermal growth factor receptor or	
PT	analogues - and comprise at least two different cell surface binding	
XX	domain(s), useful for tumour therapy.	
PS	Example 10, Page 25-27; 52pp; English.	
XX		
XX	scFv(FRP5)/TGF α -ETA (AAW05138) comprises the single-chain binding	
CC	region of murine monoclonal antibody FRP5 (specific for human epidermal	
CC	growth factor receptor erbB-2, see also AAW05134) joined to portions of	
CC	exotoxin A from Pseudomonas aeruginosa and to human transforming growth	
CC	factor (TGF) α (see also AAW05137). It is encoded by plasmid pMS238-5	
CC	-TGF (AAT42038). This plasmid can be utilised in the prodn. of the	
CC	bispecific fusion protein in bacterial (esp. E. coli) host cells. Such	
XX	fusion proteins (see also AAW05139-44) are useful as antitumour agents	
XX		
SQ	Sequence 699 AA;	
	Query Match 55.5%; Score 1302; DB 2; Length 699;	
	Best Local Similarity 99.2%; Pred. No. 1.5e-82;	
	Matches 242; Conservative 1; Mismatches 1; Indels 0; Gaps 0;	
QY	1 QIQLVQSGPELKKPEETVYKISKASGYPTNTGMMVWKOAPQGLKMMGWINSTSGESTF 60	
	..	
DB	39 QVQLQDSGPPELKKPEETVYKISKASGYPTNTGMMVWKOAPQGLKMMGWINSTSGESTF 98	
QY	61 ADDFKGRDFDSLETSAANTAYLQINNLSKEDMATYCARREYVHGVPYWGQTTVTYVSSG 120	
DB	99 ADDFKGRDFDSLETSAANTAYLQINNLSKEDMATYCARREYVHGVPYWGQTTVTYVSSG 158	
QY	121 GGGSGGGSGGGGSDIQLTQSHKFLSTSVGDRVSIITCKASQDVYNAVAVMYQKPGQSPKL 180	
DB	159 GGGSGGGSGGGGSDIQLTQSHKFLSTSVGDRVSIITCKASQDVYNAVAVMYQKPGQSPKL 218	
QY	181 LIYSASSRYTVGPSRFTSGSGSPDPTFTITSSVQADDLAVYFCQGHFRPTFTGSGTKLEI 240	
DB	219 LIYSASSRYTVGPSRFTSGSGSPDPTFTITSSVQADDLAVYFCQGHFRPTFTGSGTKLEI 278	
QY	241 KALE 244	
DB	279 KALE 282	
	RESULT 11	
XX	AAW05140	
AC	AAW05140 standard; protein; 892 AA.	
XX		
XX	AAW05140;	
XX		
DT	29-JAN-1997 (first entry)	
XX		

DE	bcrFv2(225/FRP5)-ETA.
XX	Single chain antibody; scFv; monoclonal antibody; Mab; EGF; erbB-2;
KW	epidermal growth factor; receptor; plasmid pMS238-225-5; cancer;
XX	exotoxin A; ETn; antitumour.
OS	Mus; sp.
OS	Pseudomonas; aeruginosa.
OS	Synthetic.
OS	Chimeric.
FH	
FT	Key
FT	Peptide
FT	/location/Qualifiers
FT	1..21
FT	/label= Sig_peptide
FT	/note= "ompa signal peptide"
FT	22..38
FT	/label= Spacer
FT	39..278
FT	/label= bcfV(225)
FT	279..289
FT	/label= Spacer
FT	290..404
FT	/label= ETA
FT	/note= "exotoxin A amino acids 252-366"
FT	405..407
FT	/label= Spacer
FT	408..647
FT	/label= scfv(FRP5)
FT	648..658
FT	/label= Spacer
FT	659..892
FT	/label= ETA
FT	/note= "endotoxin-A amino acids 380-613"
XX	
PN	EP739984-A1.
PD	30-OCT-1996.
XX	
PF	26-APR-1995; 9SEP-00106275.
PR	26-APR-1995; 9SEP-00106275.
PA	(SANT-) SAN TUMORFORSCHUNGS GMBH.
XI	
PI	Weis W, Schmidt M, Groner B,
DR	MPJ; 1996-478748/48.
N-PSDB;	AAT42040.
Bivalent fusion proteins that bind epidermal growth factor receptor or analogues - and comprise at least two different cell surface binding domain(s), useful for tumour therapy.	
Example 11; Page 31-33; 52pp; English.	
sbcFv2(225/FRP5)-ETA (AAW05140) comprises the single-chain binding region of murine monoclonal antibody 225 (specific for human epidermal growth factor receptor, see also AAW05133) joined to portions of exotoxin A from Pseudomonas aeruginosa and to the single-chain binding region of murine monoclonal antibody FRP5 (specific for human epidermal growth factor receptor erbB-2, see also AAW05134). It is encoded by plasmid pMS238-225-5 (AAT42040). This plasmid can be utilised in the prodn. of the bivalent fusion protein in bacterial (esp. E. coli) host cells. Such fusion proteins (see also AAW05138-44) are useful as antitumour agents	
Sequence 892 AA:	
Query Match	55.5%; Score 1302; DB 2; Length 892;
Best Local Similarity	99.2%; Pred. No. 2e-82;
Matches 242; Conservative	1; Mismatches 1; Indels 0; Gaps 0;
1 QIQLVGGPELKKKGETIVKISCKASGYPTNYGMMWKQPGGLXNMGMINTSGESTF 60	
:	

Db 408 QVQLQSGPELTKKPGETVKISCKASGYPTNYGMNWKQAPGQGLKMMGMINTSTGESTF 467
 QY 61 ADPKGRFDFSLSTANTAVYLIQINNLSKSEDMATYFCARMEVYHGVPYWGQGTITVYSSG 120
 Db 468 ADPKGRFDFSLSTANTAVYLIQINNLSKSEDMATYFCARMEVYHGVPYWGQGTITVYSSG 527
 QY 121 GGGSGGGSGGGSDIQLTQSHKFLSTSVGDRVSITCKASQDVYNAVAMVQKPGQSPKL 180
 Db 528 GGGSGGGSGGGSDIQLTQSHKFLSTSVGDRVSITCKASQDVYNAVAMVQKPGQSPKL 587
 QY 181 LIYSASSRYTGVSPRFTGSGSGDPFTFTISSVQAEADLAVYFCQGHFRTPTFGSGTKLEI 240
 Db 588 LIYSASSRYTGVSPRFTGSGSGDPFTFTISSVQAEADLAVYFCQGHFRTPTFGSGTKLEI 647
 QY 241 KALE 244
 Db 648 KALE 651

RESULT 12

AAW05143
 ID AAW05143 standard; protein; 892 AA.

XX AC AAW05143;
 XX 29-JAN-1997 (first entry)
 DE scFv2 (FRP5/FRP5)-ETA (version 2).

XX Single chain antibody; scFv; monoclonal antibody; Mab; EGF; erbB-2;
 KM epidermal growth factor; receptor; plasmid pMS238-5-5; cancer;
 KM exotoxin A; ETA; antitumour.

XX Mus; sp.
 OS Pseudomonas; aeruginosa.
 OS Synthetic.
 OS Chimeric.

XX Key Location/Qualifiers

FT Peptide 1..21

FT Peptide /label= Sig_peptide

FT Peptide /note= "ompa signal peptide"

FT Peptide 22..38

FT Peptide /label= Spacer

FT Peptide 39..278

FT Peptide /label= scFv(FRP5)

FT Peptide 279..289

FT Peptide /label= Spacer

FT Peptide 290..404

FT Peptide /label= ETA

FT Peptide /note= "exotoxin A amino acids 252-366"

FT Peptide 405..407

FT Peptide /label= Spacer

FT Peptide 408..647

FT Peptide /label= scFv(FRP5)

FT Peptide 648..658

FT Peptide /label= Spacer

FT Peptide 659..892

FT Peptide /label= ETA

FT Peptide /note= "exotoxin A amino acids 380-613"

DR WPI; 1996-4/78748/48.
 DR N-PSDB; AAT42043.
 XX Bivalent fusion proteins that bind epidermal growth factor receptor or
 FT analogues - and comprise at least two different cell surface binding
 PT domain(s), useful for tumour therapy.
 XX Example 12; Page 40-42; 52pp; English.
 CC scFv2 (FRP5/FRP5)-ETA version 2 (AAW05143) includes 2 copies of the single
 CC -chain binding region of murine monoclonal antibody FRP5 (specific for
 CC human epidermal growth factor receptor erbB-2, see also AAW05144) joined
 CC to portions of exotoxin A from Pseudomonas aeruginosa. It is encoded by
 CC plasmid pMS238-5-5 (AAT42043). This plasmid can be utilised in the prodn.
 CC of the bivalent fusion protein in bacterial (esp. E. coli) host cells.
 CC Such fusion proteins (see also AAW05138-44) are useful as antitumour
 CC agents
 CC Sequence 892 AA;
 SQ

Query Match 55.5%; Score 1302; DB 2; Length 892;
 Best Local Similarity 99.2%; Pred. No. 2e-82; Indels 0; Gaps 0;
 Matches 242; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QIQLVQSGPELTKKPGETVKISCKASGYPTNYGMNWKQAPGQGLKMMGMINTSTGESTF 60

Db 39 QVQLQSGPELTKKPGETVKISCKASGYPTNYGMNWKQAPGQGLKMMGMINTSTGESTF 98

QY 61 ADPKGRFDFSLSTANTAVYLIQINNLSKSEDMATYFCARMEVYHGVPYWGQGTITVYSSG 120

Db 99 ADPKGRFDFSLSTANTAVYLIQINNLSKSEDMATYFCARMEVYHGVPYWGQGTITVYSSG 158

QY 121 GGGSGGGSGGGSDIQLTQSHKFLSTSVGDRVSITCKASQDVYNAVAMVQKPGQSPKL 180

Db 159 GGGSGGGSGGGSDIQLTQSHKFLSTSVGDRVSITCKASQDVYNAVAMVQKPGQSPKL 218

QY 181 LIYSASSRYTGVSPRFTGSGSGDPFTFTISSVQAEADLAVYFCQGHFRTPTFGSGTKLEI 240

Db 219 LIYSASSRYTGVSPRFTGSGSGDPFTFTISSVQAEADLAVYFCQGHFRTPTFGSGTKLEI 278

QY 241 KALE 244

Db 279 KALE 282

RESULT 13

AAW05139
 ID AAW05139 standard; protein; 892 AA.

XX AC AAW05139;
 XX 29-JAN-1997 (first entry)

DE scFv2 (FRP5/225)-ETA (version 1).

XX Single chain antibody; scFv; monoclonal antibody; Mab; EGF; erbB-2;
 KM epidermal growth factor; receptor; plasmid pMS238-5-225; cancer;
 KM exotoxin A; ETA; antitumour.

XX Mus; sp.
 OS Pseudomonas; aeruginosa.
 OS Synthetic.
 OS Chimeric.

XX Key Location/Qualifiers

FT Peptide 1..21

FT Peptide /label= Sig_peptide

FT Peptide /note= "ompa signal peptide"

FT Peptide 22..38

FT Peptide /label= Spacer

FT Peptide 39..278

FT Peptide /label= scFv(FRP5)

FT Peptide 279..289

```

FT FT /label= Spacer
FT FT 290..404
FT FT /label= ETA
FT FT /note= "exotoxin A amino acids 252-366"
FT FT Peptide
FT FT 405..407
FT FT /label= Spacer
FT FT Region
FT FT 408..647
FT FT /label= scFv(225)
FT FT Peptide
FT FT 648..658
FT FT /label= Spacer
FT FT Region
FT FT 659..892
FT FT /label= ETA
FT FT /note= "endotoxin-A amino acids 380-613"

XX XX EP739984-A1.
XX XX 30-OCT-1996.
XX XX 26-APR-1995; 95EP-00106275.
XX XX 26-APR-1995; 95EP-00106275.
XX XX (SANT-) SAN TUMORFORSCHUNGS GMBH.
XX XX Wels W, Schmidt M, Groner B;
XX XX WPI; 1996-478748/48.
XX XX N-PSDB; AAT42039.
XX XX Bivalent fusion proteins that bind epidermal growth factor receptor or
XX XX analogues - and comprise at least two different cell surface binding
XX XX domain(s), useful for tumour therapy.
XX XX Example 11; Page 28-30; 52pp; English.
XX XX scFv2(FRP5/225)-ETA (AAW05139) comprises the single-chain binding region
XX XX of murine monoclonal antibody FRP5 (specific for human epidermal growth
XX XX factor receptor erbB-2, see also AAW05134) joined to portions of exotoxin
XX XX A from Pseudomonas aeruginosa and to the single-chain binding region of
XX XX murine monoclonal antibody 225 (specific for human epidermal growth
XX XX factor receptor, see also AAW05133). It is encoded by plasmid pMS238-5-
XX XX 225 (AAT42039). This plasmid can be utilised in the prodn. of the
XX XX bivalent fusion protein in bacterial (esp. E. coli) host cells. Such
XX XX fusion proteins (see also AAW05138-44) are useful as antitumour agents
XX XX
XX XX Sequence 892 AA:
XX XX
XX XX Query Match 55.5%; Score 1302; DB 2; Length 892;
XX XX Best Local Similarity 99.2%; Pred. No. 2e-82;
XX XX Matches 242; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX XX
XX XX 1 QIQVQSGPELKKPEPTKISCKASGYPTNYGMWVQAQPGQIKMGMINTSTGESETF 60
XX XX :|||
XX XX 39 QVQLQDSGPPELKKPEPTKISCKASGYPTNYGMWVQAQPGQIKMGMINTSTGESETF 98
XX XX :|||
XX XX 61 ADPFKGRDFSLSTANTAYLQINNLSKSDMATYFCAMENVHGYVPWGGTTVTYSSG 120
XX XX :|||
XX XX 99 ADPFKGRDFSLSTANTAYLQINNLSKSDMATYFCAMENVHGYVPWGGTTVTYSSG 158
XX XX :|||
XX XX 121 GGGSGGGSGGGSDIQLTQSHKFLSTSVGDVRSITCKASQDVYNAVAMYOQKPGQSPKL 180
XX XX :|||
XX XX 159 GGGSGGGSGGGSDIQLTQSHKFLSTSVGDVRSITCKASQDVYNAVAMYOQKPGQSPKL 218
XX XX :|||
XX XX 181 LIYSASSRYTVPSRFTGSGGPDFTTSSVQAEDLAVYPCQGHFRTPTFGSGTKLEI 240
XX XX :|||
XX XX 219 LIYSASSRYTVPSRFTGSGGPDFTTSSVQAEDLAVYPCQGHFRTPTFGSGTKLEI 278
XX XX :|||
XX XX 241 KALE 244
XX XX :|||
XX XX 279 KALE 282
XX XX
XX XX RESULT 14

```

```

AAW05142
XX ID AAW05142 standard; protein; 895 AA.
XX AC AAW05142;
XX XX 29-JAN-1997 (first entry)
XX XX scFv2(FRP5/FRP5)-ETA (version 1).
XX XX DE scFv2(FRP5/FRP5)-ETA (version 1).
XX XX KW Single chain antibody; scFv; monoclonal antibody; Mab; EGF; erbB-2;
XX XX KM epidermal growth factor; receptor; plasmid pMS242-5-5; cancer;
XX XX KW exotoxin A; ETA; antitumour.
XX XX
XX XX Mus; sp.
XX XX OS Pseudomonas; aeruginosa.
XX XX OS Synthetic.
XX XX OS Chimeric.
XX XX
XX XX Key
XX XX Location/Qualifiers
XX XX FT 1..21
XX XX FT /label= Sig peptide
XX XX FT /note= "ompA signal peptide"
XX XX FT Peptide
XX XX 22..44
XX XX FT /label= Spacer
XX XX FT 45..159
XX XX FT /label= ETA
XX XX FT /note= "exotoxin A amino acids 252-366"
XX XX FT 160..162
XX XX FT /label= Spacer
XX XX FT 163..402
XX XX FT /label= scFv(FRP5)
XX XX FT Peptide
XX XX 403..413
XX XX FT /label= Spacer
XX XX FT 414..644
XX XX FT /label= ETA
XX XX FT /note= "exotoxin A amino acids 380-610"
XX XX FT Peptide
XX XX 645..646
XX XX FT /label= Spacer
XX XX FT 647..886
XX XX FT /label= scFv(FRP5)
XX XX FT Peptide
XX XX 887..895
XX XX FT /label= Spacer
XX XX
XX XX EP739984-A1.
XX XX 30-OCT-1996.
XX XX 26-APR-1995; 95EP-00106275.
XX XX 26-APR-1995; 95EP-00106275.
XX XX 26-APR-1995; 95EP-00106275.
XX XX (SANT-) SAN TUMORFORSCHUNGS GMBH.
XX XX Wels W, Schmidt M, Groner B;
XX XX WPI; 1996-478748/48.
XX XX N-PSDB; AAT42042.
XX XX
XX XX Bivalent fusion proteins that bind epidermal growth factor receptor or
XX XX analogues - and comprise at least two different cell surface binding
XX XX domain(s), useful for tumour therapy.
XX XX Example 12; Page 37-39; 52pp; English.
XX XX
XX XX scFv2(FRP5/FRP5)-ETA version 1 (AAW05142) includes 2 copies of the single
XX XX -chain binding region of murine monoclonal antibody FRP5 (specific for
XX XX human epidermal growth factor receptor erbB-2, see also AAW05134) joined
XX XX to portions of exotoxin A from Pseudomonas aeruginosa. It is encoded by
XX XX plasmid pMS242-5-5 (AAT42042). This plasmid can be utilised in the prodn.
XX XX of the bivalent fusion protein in bacterial (esp. E. coli) host cells.
XX XX Such fusion proteins (see also AAW05138-44) are useful as antitumour
XX XX agents
XX XX

```

SQ Sequence 895 AA;
 Query Match 55.5%; Score 1302; DB 2; Length 895;
 Best Local Similarity 99.2%; Pred. No. 2e-82;
 Matches 242; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 QIQLVSGPELKKRGETVTKISCKASGYPTNYGNMVKQAPQGLKMMGMINSTGSESTF 60
 DB 163 QVQLQDSGPELKKRGETVTKISCKASGYPTNYGNMVKQAPQGLKMMGMINSTGSESTF 222
 QY 61 ADDPKGRFDFSLFETSAANTAYLQINNLKSEDMATYFCARMEVYHGVPYWGQGTITVYSSG 120
 DB 223 ADDPKGRFDFSLFETSAANTAYLQINNLKSEDMATYFCARMEVYHGVPYWGQGTITVYSSG 282
 QY 121 GGSAGSGSGSGGSDIQLTQSHKFLSTSVGDRVSITCKASQDYNNAVAMVQOKPGQSPKL 180
 DB 283 GGSAGSGSGSGGSDIQLTQSHKFLSTSVGDRVSITCKASQDYNNAVAMVQOKPGQSPKL 342
 QY 181 LIYSASSRYTGVPSRFTGSGSGPDFTFTISSVQAEIDLAVYFCQGHFRTPTFGSGTKLEI 240
 DB 343 LIYSASSRYTGVPSRFTGSGSGPDFTFTISSVQAEIDLAVYFCQGHFRTPTFGSGTKLEI 402
 QY 241 KALE 244
 DB 403 KALE 406
 RESULT 15
 AAM05144
 ID AAM05144 standard; protein; 899 AA.
 AC AAM05144;
 XX 29-JAN-1997 (first entry)
 DE scFv2(FRP5/FRP5)-ETA (version 3).
 XX Single chain antibody; scFv; monoclonal antibody; Mab; EGF; erbB-2;
 KM epidermal growth factor; receptor; plasmid pMS246-5-5; cancer;
 KM exotoxin A; ETA; antitumour.
 XX Mus; sp.
 OS Pseudomonas; aeruginosa.
 OS Synthetic.
 OS Chimeric.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..21
 FT /label= Sig_peptide
 FT /note= "ompa signal peptide"
 FT 22..38
 FT /label= Spacer
 FT Region 39..278
 FT /label= scFv (FRP5)
 FT Peptide 279..289
 FT /label= Spacer
 FT Region 290..648
 FT /label= ETA
 FT /note= "exotoxin A amino acids 252-610"
 FT Peptide 649..650
 FT /label= Spacer
 FT Region 651..890
 FT /label= scFv (FRP5)
 FT Peptide 891..899
 FT /label= Spacer
 XX
 XX EP73984-A1.
 XX
 PD 30-OCT-1996.
 XX
 XX 26-APR-1995; 95EP-00106275.
 XX 26-APR-1995; 95EP-00106275.
 PR

XX (SANT-) SAN TUMORFORSCHUNGS GMBH.
 PA
 XX
 XX Wels W, Schmidt M, Groner B;
 PI
 XX
 XX WPI: 1996-478748/48.
 DR
 DR N-PSDB; AAT42044.
 XX
 XX Bivalent fusion proteins that bind epidermal growth factor receptor or
 PT analogues - and comprise at least two different cell surface binding
 PT domain(s), useful for tumour therapy.
 XX
 XX Example 12; Page 43-45; 52pp; English.
 PS
 CC scFv2(FRP5/FRP5)-ETA version 3 (AAM05144) includes 2 copies of the single
 CC chain binding region of murine monoclonal antibody FRP5 (specific for
 CC human epidermal growth factor receptor erbB-2, see also AAM05134) joined
 CC to portions of exotoxin A from Pseudomonas aeruginosa. It is encoded by
 CC plasmid pMS246-5-5 (AAT42044). This plasmid can be utilised in the prodn.
 CC of the bivalent fusion protein in bacterial (esp. E. coli) host cells.
 CC Such fusion proteins (see also AAM05138-43) are useful as antitumour
 CC agents
 CC
 XX
 XX Sequence 899 AA;
 SQ
 Query Match 55.5%; Score 1302; DB 2; Length 899;
 Best Local Similarity 99.2%; Pred. No. 2e-82;
 Matches 242; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 QIQLVSGPELKKRGETVTKISCKASGYPTNYGNMVKQAPQGLKMMGMINSTGSESTF 60
 DB 39 QVQLQDSGPELKKRGETVTKISCKASGYPTNYGNMVKQAPQGLKMMGMINSTGSESTF 98
 QY 61 ADDPKGRFDFSLFETSAANTAYLQINNLKSEDMATYFCARMEVYHGVPYWGQGTITVYSSG 120
 DB 99 ADDPKGRFDFSLFETSAANTAYLQINNLKSEDMATYFCARMEVYHGVPYWGQGTITVYSSG 158
 QY 121 GGSAGSGSGSGGSDIQLTQSHKFLSTSVGDRVSITCKASQDYNNAVAMVQOKPGQSPKL 180
 DB 159 GGSAGSGSGSGGSDIQLTQSHKFLSTSVGDRVSITCKASQDYNNAVAMVQOKPGQSPKL 218
 QY 181 LIYSASSRYTGVPSRFTGSGSGPDFTFTISSVQAEIDLAVYFCQGHFRTPTFGSGTKLEI 240
 DB 219 LIYSASSRYTGVPSRFTGSGSGPDFTFTISSVQAEIDLAVYFCQGHFRTPTFGSGTKLEI 278
 QY 241 KALE 244
 DB 279 KALE 282

Search completed: April 20, 2005, 07:04:53
 Job time : 66.1759 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2005, 07:01:08 ; Search time 20.0918 Seconds
(without alignments)
1645.921 Million cell updates/sec

Title: US-09-596-774-7
Perfect score: 2345
Sequence: 1 QIOLVSGPELKKKGFTYK1.....LSTATKDTYDALHMQTLAPR 443

Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1308	55.8	615	4	US-08-840-713-35
2	1308	55.8	617	4	US-08-840-713-37
3	1303	55.6	711	1	US-08-235-838-7
4	1303	55.6	711	2	US-08-465-473B-7
5	1302	55.5	530	4	US-08-840-713-2
6	1302	55.5	637	1	US-08-235-838-14
7	1302	55.5	637	2	US-08-465-473B-14
8	1284	54.8	241	2	US-08-235-838-5
9	1284	54.8	241	2	US-08-465-473B-5
10	981.5	41.9	250	1	US-08-133-804-2
11	981.5	41.9	250	1	US-08-461-184-8
12	981.5	41.9	250	1	US-08-463-675-8
13	981.5	41.9	250	1	US-08-464-589-8
14	981.5	41.9	250	1	US-08-461-838-2
15	981.5	41.9	250	1	US-08-461-386-2
16	975.5	41.6	622	2	US-08-356-786-16
17	903.5	38.5	240	1	US-08-488-113B-148
18	903.5	38.5	240	1	US-08-477-484B-148
19	903.5	38.5	240	2	US-08-646-360-148
20	903.5	38.5	240	3	US-08-839-765-148
21	903.5	38.5	240	3	US-09-136-389-148
22	903.5	38.5	240	3	US-09-610-838-148
23	903.5	38.5	240	4	US-09-711-485-148
24	891	38.0	366	3	US-08-875-811-55
25	890	38.0	365	3	US-08-875-811-53
26	889.5	37.9	541	3	US-09-485-737B-85
27	889.5	37.9	541	4	US-10-071-485-85

28	888.5	37.9	267	3	US-09-485-737B-2	Sequence 2, Appli
29	888.5	37.9	267	4	US-10-071-485-2	Sequence 2, Appli
30	888.5	37.9	711	3	US-09-485-737B-90	Sequence 90, Appli
31	888.5	37.9	711	4	US-10-071-485-90	Sequence 90, Appli
32	885.5	37.8	259	4	US-09-419-788-115	Sequence 115, App
33	884.5	37.7	284	3	US-09-184-658-40	Sequence 40, Appli
34	884.5	37.7	284	4	US-09-504-262D-40	Sequence 40, Appli
35	878.5	37.5	534	2	US-08-356-786-10	Sequence 10, Appli
36	851	36.3	240	4	US-10-092-246-35	Sequence 35, Appli
37	851	36.3	240	4	US-10-092-246-36	Sequence 36, Appli
38	851	36.3	240	4	US-10-096-246A-35	Sequence 35, Appli
39	851	36.3	240	4	US-10-096-246A-37	Sequence 37, Appli
40	848	36.2	240	3	US-09-485-737B-91	Sequence 91, Appli
41	848	36.2	240	4	US-10-071-485-91	Sequence 91, Appli
42	843	35.9	240	4	US-10-096-246A-36	Sequence 36, Appli
43	841.5	35.9	249	2	US-08-797-689-18	Sequence 18, Appli
44	841.5	35.9	249	4	US-09-984-186-18	Sequence 18, Appli
45	839	35.8	240	4	US-10-092-246-37	Sequence 37, Appli

ALIGNMENTS

RESULT 1
US-08-840-713-35
Sequence 35, Application US/08840713
Patent No. 6498233
GENERAL INFORMATION:
APPLICANT: WELS, Winfried, Dr.
APPLICANT: FOYMINAYA, Jesus
TITLE OF INVENTION: NUCLEIC ACID TRANSFER SYSTEM
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nikolaic, Marmelstein, Murray & Oram LLP
STREET: 655 15th St., N.W., Suite 330 - G St. Lobby
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-5701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/840,713
FILING DATE: 25-APR-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Kites, Monica Chin
REGISTRATION NUMBER: 36,105
REFERENCE/DOCKET NUMBER: 1614-7014
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 638 - 5000
TELEFAX: (202) 638 - 4810
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 615 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-840-713-35
Query Match 55.8%; Score 1308; DB 4; Length 615;
Best Local Similarity 96.5%; Pred. No. 4.3e-100;
Matches 245; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
QY 1 QIOLVSGPELKKKGFTYK1SCASGYPTNYGNWYKQAPGGLKMMGWINTSGSTF 60
DB 362 QVQLQSGPELKKKGFTYK1SCASGYPTNYGNWYKQAPGGLKMMGWINTSGSTF 421
QY 61 ADKFGKDFPSLETSTANTAYLIQINLKSSEDATYFCARWEVYHGVYPMGGITVTVSSG 120

Db 422 ADDEKGRFDFSLSTANTAYLIQINNLKSEDMATYFCARMEVYHGVPYWGQTTVTVSSG 481
QY 121 GGGSGGGSGGGSDIQLTQSHKFLSTSVGRVSTTCASQDVYNAVAMYOQKPGQSPKL 180
Db 482 GGGSGGGSGGGSDIQLTQSHKFLSTSVGRVSTTCASQDVYNAVAMYOQKPGQSPKL 541
QY 181 LIYSASRYTGVPSRFTGSGGPDFTFTISSVQAEIDLAVYFCQGHFRPTFTFGSGTKLEI 240
Db 542 LIYSASRYTGVPSRFTGSGGPDFTFTISSVQAEIDLAVYFCQGHFRPTFTFGSGTKLEI 601
QY 241 KALEISNSVMFSS 254
Db 602 KALEDLSSERRFSA 615

RESULT 2

US-08-840-713-37
Sequence 37, Application US/08840713
Patent No. 6498233
GENERAL INFORMATION:
APPLICANT: WELLS, Winfried, Dr.
APPLICANT: FOYMINAYA, Jesus
TITLE OF INVENTION: NUCLEIC ACID TRANSFER SYSTEM
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nikaido, Marmelestein, Murray & Oram LLP
STREET: 655 15th St., N.W., Suite 330 - G St. Lobby
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-5701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/840,713
FILING DATE: 25-APR-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Kiltz, Monica Chin
REGISTRATION NUMBER: 36,105
REFERENCE/DOCKET NUMBER: 1614-7014
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 638 - 5000
TELEFAX: (202) 638 - 4810
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 617 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-840-713-37

Query Match 55.8%; Score 1308; DB 4; Length 617;
Best Local Similarity 96.5%; Pred. No. 4.3e-100;

Matches 245; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 QIQLVSGPELKKRGFTVTKISCKASGYPTNYGNMVKQAPQGLKMMGMINTSGESTF 60
Db 364 QVQLQSGPELKKRGFTVTKISCKASGYPTNYGNMVKQAPQGLKMMGMINTSGESTF 423
QY 61 ADDEKGRFDFSLSTANTAYLIQINNLKSEDMATYFCARMEVYHGVPYWGQTTVTVSSG 120
Db 424 ADDEKGRFDFSLSTANTAYLIQINNLKSEDMATYFCARMEVYHGVPYWGQTTVTVSSG 483
QY 121 GGGSGGGSGGGSDIQLTQSHKFLSTSVGRVSTTCASQDVYNAVAMYOQKPGQSPKL 180
Db 484 GGGSGGGSGGGSDIQLTQSHKFLSTSVGRVSTTCASQDVYNAVAMYOQKPGQSPKL 543
QY 181 LIYSASRYTGVPSRFTGSGGPDFTFTISSVQAEIDLAVYFCQGHFRPTFTFGSGTKLEI 240

Db 544 LIYSASRYTGVPSRFTGSGGPDFTFTISSVQAEIDLAVYFCQGHFRPTFTFGSGTKLEI 603
QY 241 KALEISNSVMFSS 254
Db 604 KALEDLSSERRFSA 617

RESULT 3

US-08-235-838-7
Sequence 7, Application US/08235838
Patent No. 5571894
GENERAL INFORMATION:
APPLICANT: Wells, Winfried S.
APPLICANT: Hynes, Nancy E.
APPLICANT: Harwerth, Ina-Maria
APPLICANT: Groner, Bernd
APPLICANT: Hardman, No. 5571894man
APPLICANT: Zwickl, Markus
TITLE OF INVENTION: Recombinant Antibodies Specific for a
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: New York
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235,838
FILING DATE: TBA
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/828,832
FILING DATE: 31-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 91-810079.3
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: 4-18518/A/CIP/CONT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8614
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 711 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-235-838-7

Query Match 55.6%; Score 1303; DB 1; Length 711;
Best Local Similarity 94.2%; Pred. No. 1.4e-99;

Matches 245; Conservative 2; Mismatches 1; Indels 12; Gaps 1;

QY 1 QIQLVSGPELKKRGFTVTKISCKASGYPTNYGNMVKQAPQGLKMMGMINTSGESTF 60
Db 23 QVQLQSGPELKKRGFTVTKISCKASGYPTNYGNMVKQAPQGLKMMGMINTSGESTF 82
QY 61 ADDEKGRFDFSLSTANTAYLIQINNLKSEDMATYFCARMEVYHGVPYWGQTTVTVSSG 120
Db 83 ADDEKGRFDFSLSTANTAYLIQINNLKSEDMATYFCARMEVYHGVPYWGQTTVTVSSG 142
QY 121 GGGSGGGSGGGSDIQLTQSHKFLSTSVGRVSTTCASQDVYNAVAMYOQKPGQSPKL 180
Db 143 GGGSGGGSGGGSDIQLTQSHKFLSTSVGRVSTTCASQDVYNAVAMYOQKPGQSPKL 202

Qy 181 LIYSASSRYTGVPSRFTSGSGPDFTFTISSVOAEDLAVYFCQOHFRPTFTGSGTKLEI 240
Db 203 LIYSASSRYTGVPSRFTSGSGPDFTFTISSVOAEDLAVYFCQOHFRPTFTGSGTKLEI 262
Qy 241 KALEISNSVMYFSSVVPYLO 260
Db 263 KALE-----PVLE 270

RESULT 4

US-08-465-473B-7
; Sequence 7, Application US/08465473B
; Patent No. 5939531
; GENERAL INFORMATION:
; APPLICANT: Wels, Winfried S.
; APPLICANT: Hynes, Nancy E.
; APPLICANT: Harwerth, Ina-Maria
; APPLICANT: Groner, Bernd
; APPLICANT: Hardman, No. 5939531man
; APPLICANT: Zwickl, Markus
; TITLE OF INVENTION: Recombinant Antibodies Specific for a
; GROWTH FACTOR RECEPTOR
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NOVARTIS Corporation
; STREET: 564 Morris Avenue
; CITY: Summit
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07901-6940
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,473B
; FILING DATE: 5 June 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/828,832
; FILING DATE: 31-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 91-810079.3
; FILING DATE: 05-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Pfeiffer, Heena J.
; REGISTRATION NUMBER: 22,640
; REFERENCE/DOCKET NUMBER: 4-18518/A/CIP/CONT2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908)522 6940
; TELEFAX: (908)522 6955
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 711 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-465-473B-7

Query Match 55.6%; Score 1303; DB 2; Length 711;
Best Local Similarity 94.2%; Pred. No. 1.4e-99;
Matches 245; Conservative 2; Mismatches 1; Indels 12; Gaps 1;
Qy 1 QIOLVQSGPELKKPGETYKISCKASGYPTNYGMNMYQAQPGQGLKMMGMINTSGESTF 60
Db 23 QVQLQDSGPPELKKPGETYKISCKASGYPTNYGMNMYQAQPGQGLKMMGMINTSGESTF 82
Qy 61 ADDPKGRDFSLFETISANTAYLIQINNLSKEDMATYFCARMEVYHGVVPYWGQTTVTVSSG 120
Db 83 ADDPKGRDFSLFETISANTAYLIQINNLSKEDMATYFCARMEVYHGVVPYWGQTTVTVSSG 142

Qy 121 GGSGGGGGGGGGSDIQLTQSHKFLSTVGDVRSITCKASQDVYNAVAYQKPGQSPKL 180
Db 143 GGSGGGGGGGGGSDIQLTQSHKFLSTVGDVRSITCKASQDVYNAVAYQKPGQSPKL 202
Qy 181 LIYSASSRYTGVPSRFTSGSGPDFTFTISSVOAEDLAVYFCQOHFRPTFTGSGTKLEI 240
Db 203 LIYSASSRYTGVPSRFTSGSGPDFTFTISSVOAEDLAVYFCQOHFRPTFTGSGTKLEI 262
Qy 241 KALEISNSVMYFSSVVPYLO 260
Db 263 KALE-----PVLE 270

RESULT 5

US-08-840-713-2
; Sequence 2, Application US/08840713
; Patent No. 6498223
; GENERAL INFORMATION:
; APPLICANT: WELS, Winfried, Dr.
; APPLICANT: FOYMINAYA, Jesus
; TITLE OF INVENTION: NUCLEIC ACID TRANSFER SYSTEM
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP
; STREET: 655 15th St., N.W., Suite 330 - G St. Lobby
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/840,713
; FILING DATE: 25-APR-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Kites, Monica Chiu
; REGISTRATION NUMBER: 36,105
; REFERENCE/DOCKET NUMBER: 1614-7014
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 638 - 5000
; TELEFAX: (202) 638 - 4810
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-840-713-2

Query Match 55.5%; Score 1302; DB 4; Length 530;
Best Local Similarity 99.2%; Pred. No. 1.1e-99;
Matches 242; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 1 QIOLVQSGPELKKPGETYKISCKASGYPTNYGMNMYQAQPGQGLKMMGMINTSGESTF 60
Db 18 QVQLQDSGPPELKKPGETYKISCKASGYPTNYGMNMYQAQPGQGLKMMGMINTSGESTF 77
Qy 61 ADDPKGRDFSLFETISANTAYLIQINNLSKEDMATYFCARMEVYHGVVPYWGQTTVTVSSG 120
Db 78 ADDPKGRDFSLFETISANTAYLIQINNLSKEDMATYFCARMEVYHGVVPYWGQTTVTVSSG 137
Qy 121 GGSGGGGGGGGGSDIQLTQSHKFLSTVGDVRSITCKASQDVYNAVAYQKPGQSPKL 180
Db 138 GGSGGGGGGGGGSDIQLTQSHKFLSTVGDVRSITCKASQDVYNAVAYQKPGQSPKL 197
Qy 181 LIYSASSRYTGVPSRFTSGSGPDFTFTISSVOAEDLAVYFCQOHFRPTFTGSGTKLEI 240
Db 198 LIYSASSRYTGVPSRFTSGSGPDFTFTISSVOAEDLAVYFCQOHFRPTFTGSGTKLEI 257

QY 241 KALE 244
|||
Db 258 KALE 261

RESULT 6
US-08-235-838-14

; Sequence 14, Application US/08235838
; Patent No. 5571894

GENERAL INFORMATION:

APPLICANT: Wels, Winfried S.
APPLICANT: Hynes, Nancy E.
APPLICANT: Harwerth, Ina-Maria
APPLICANT: Groner, Bernd
APPLICANT: Hardman, No. 5571894man
APPLICANT: Zwickl, Markus
TITLE OF INVENTION: Recombinant Antibodies Specific for a
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESSES:
ADDRESSEE: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: New York
COUNTRY: USA
ZIP: 10532

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235, 838
FILING DATE: TBA

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/828, 832
FILING DATE: 31-JAN-1992

PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 91-810079.3
FILING DATE: 05-FEB-1991

ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129

REFERENCE/DOCKET NUMBER: 4-18518/A/CIP/CONT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8614

INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 637 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-235-838-14

Query Match 55.5%; Score 1302; DB 1; Length 637;
Best Local Similarity 99.2%; Pred. No. 1.4e-99;
Matches 242; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QIQLVSGPELKKPGTIVKISCKASGYPTNYGMNWKAPQGLKMMGINTSGESTF 60
|||
Db 33 QVQLQSGPELKKPGTIVKISCKASGYPTNYGMNWKAPQGLKMMGINTSGESTF 92
|||
QY 61 ADDPKRFPDSLETSANTAYLIQINNLKSEDMATYFCARWEVHGYPYWGQSTTVVSSG 120
|||
Db 93 ADDPKRFPDSLETSANTAYLIQINNLKSEDMATYFCARWEVHGYPYWGQSTTVVSSG 152
|||
QY 121 GGGSGGGSGGGSDIQLTQSHKFLSTVGDVRSITCKASQDVYNAVAMYOQKPGQSPKL 180
|||
Db 153 GGGSGGGSGGGSDIQLTQSHKFLSTVGDVRSITCKASQDVYNAVAMYOQKPGQSPKL 212
|||
QY 181 LIYSASRYTGVSRFTGSGSGPDFTTISVQAEDLAVFCCQHRTPTFFSGTLEI 240
|||

Db 213 LIYSASRYTGVSRFTGSGSGPDFTTISVQAEDLAVFCCQHRTPTFFSGTLEI 272
|||
QY 241 KALE 244
|||
Db 273 KALE 276

RESULT 7
US-08-465-473B-14

; Sequence 14, Application US/08465473B
; Patent No. 5939531

GENERAL INFORMATION:

APPLICANT: Wels, Winfried S.
APPLICANT: Hynes, Nancy E.
APPLICANT: Harwerth, Ina-Maria
APPLICANT: Groner, Bernd
APPLICANT: Hardman, No. 5939531man
APPLICANT: Zwickl, Markus
TITLE OF INVENTION: Recombinant Antibodies Specific for a
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESSES:
ADDRESSEE: NOVARTIS Corporation
STREET: 564 Morris Avenue
CITY: Summit
STATE: New Jersey
COUNTRY: USA
ZIP: 07901-6940

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465, 473B
FILING DATE: 5 June 1995

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/828, 832
FILING DATE: 31-JAN-1992

PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 91-810079.3
FILING DATE: 05-FEB-1991

ATTORNEY/AGENT INFORMATION:
NAME: Pfeiffer, Heena J.
REGISTRATION NUMBER: 22,640

REFERENCE/DOCKET NUMBER: 4-18518/A/CIP/CONT2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908)522 6955

INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 637 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-465-473B-14

Query Match 55.5%; Score 1302; DB 2; Length 637;
Best Local Similarity 99.2%; Pred. No. 1.4e-99;
Matches 242; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QIQLVSGPELKKPGTIVKISCKASGYPTNYGMNWKAPQGLKMMGINTSGESTF 60
|||
Db 33 QVQLQSGPELKKPGTIVKISCKASGYPTNYGMNWKAPQGLKMMGINTSGESTF 92
|||
QY 61 ADDPKRFPDSLETSANTAYLIQINNLKSEDMATYFCARWEVHGYPYWGQSTTVVSSG 120
|||
Db 93 ADDPKRFPDSLETSANTAYLIQINNLKSEDMATYFCARWEVHGYPYWGQSTTVVSSG 152
|||
QY 121 GGGSGGGSGGGSDIQLTQSHKFLSTVGDVRSITCKASQDVYNAVAMYOQKPGQSPKL 180
|||

Db 153 GGGSGGGGGGGGSDIQLTOSHKFLSTSVGDVRSITCKASQDVYNAVAMYOQKPGQSPKL 212
Qy 181 LIIYASSRYTGVPRFTGSGGPDFTTISVQAEADLAVYFCQGHFRTPFTFGSGTKLEI 240
Db 213 LIIYASSRYTGVPRFTGSGGPDFTTISVQAEADLAVYFCQGHFRTPFTFGSGTKLEI 272
Qy 241 KALE 244
273 KALE 276

RESULT 8

US-08-235-838-5
; Sequence 5, Application US/08235838
; Patent No. 5571894
; GENERAL INFORMATION:
; APPLICANT: Weis, Winfried S.
; APPLICANT: Hynes, Nancy E.
; APPLICANT: Harwerth, Ina-Maria
; APPLICANT: Groner, Bernd
; APPLICANT: Hardman, No. 5571894man
; APPLICANT: Zwickl, Markus
; TITLE OF INVENTION: Recombinant Antibodies Specific for a
; TITLE OF INVENTION: Growth Factor Receptor
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: New York
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/235,838
; FILING DATE: TBA
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/828,832
; FILING DATE: 31-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 91-810079.3
; FILING DATE: 05-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Elmer, James Scott
; REGISTRATION NUMBER: 36,129
; REFERENCE/DOCKET NUMBER: 4-18518/A/CIP/CONT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8614
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 241 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-235-838-5

Query Match 54.8%; Score 1284; DB 1; Length 241;
Best Local Similarity 99.2%; Pred. No. 1.2e-98;
Matches 238; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 1 QIOLVOSGPELKKRGETIVKISCKASGYPTNYGMNWKVQAPOGGLKMMGMINSTSGESTF 60
Db 2 QVQLQDSGPELKKRGETIVKISCKASGYPTNYGMNWKVQAPOGGLKMMGMINSTSGESTF 61
Qy 61 ADDEKGRFDSLETSANTAVYLIQINNLSKSDMATYFCARMEVYHGVPYWGQGTITVTVSSG 120
Db 62 ADDEKGRFDSLETSANTAVYLIQINNLSKSDMATYFCARMEVYHGVPYWGQGTITVTVSSG 121

Qy 121 GGGSGGGGGGGGSDIQLTOSHKFLSTSVGDVRSITCKASQDVYNAVAMYOQKPGQSPKL 180
Db 122 GGGSGGGGGGGGSDIQLTOSHKFLSTSVGDVRSITCKASQDVYNAVAMYOQKPGQSPKL 181
Qy 181 LIIYASSRYTGVPRFTGSGGPDFTTISVQAEADLAVYFCQGHFRTPFTFGSGTKLEI 240
Db 182 LIIYASSRYTGVPRFTGSGGPDFTTISVQAEADLAVYFCQGHFRTPFTFGSGTKLEI 241

RESULT 9

US-08-465-473B-5
; Sequence 5, Application US/08465473B
; Patent No. 5939531
; GENERAL INFORMATION:
; APPLICANT: Weis, Winfried S.
; APPLICANT: Hynes, Nancy E.
; APPLICANT: Harwerth, Ina-Maria
; APPLICANT: Groner, Bernd
; APPLICANT: Hardman, No. 5939531man
; APPLICANT: Zwickl, Markus
; TITLE OF INVENTION: Recombinant Antibodies Specific for a
; TITLE OF INVENTION: Growth Factor Receptor
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NOVARTIS Corporation
; STREET: 564 Morris Avenue
; CITY: Summit
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07901-6940
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,473B
; FILING DATE: 5 June 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/828,832
; FILING DATE: 31-JAN-1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 91-810079.3
; FILING DATE: 05-FEB-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Pfeiffer, Heena J.
; REGISTRATION NUMBER: 22,640
; REFERENCE/DOCKET NUMBER: 4-18518/A/CIP/CONT2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (908)522 6940
; TELEFAX: (908)522 6955
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 241 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-465-473B-5

Query Match 54.8%; Score 1284; DB 2; Length 241;
Best Local Similarity 99.2%; Pred. No. 1.2e-98;
Matches 238; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
Qy 1 QIOLVOSGPELKKRGETIVKISCKASGYPTNYGMNWKVQAPOGGLKMMGMINSTSGESTF 60
Db 2 QVQLQDSGPELKKRGETIVKISCKASGYPTNYGMNWKVQAPOGGLKMMGMINSTSGESTF 61
Qy 61 ADDEKGRFDSLETSANTAVYLIQINNLSKSDMATYFCARMEVYHGVPYWGQGTITVTVSSG 120
Db 62 ADDEKGRFDSLETSANTAVYLIQINNLSKSDMATYFCARMEVYHGVPYWGQGTITVTVSSG 121

APPLICANT: Ring, David B.
TITLE OF INVENTION: Biosynthetic Binding Proteins For
TITLE OF INVENTION: Imaging
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Testa, Hurwitz & Thibault/Patent Department
STREET: Exchange Place, 53 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,838
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Kelley, Robin D.
REGISTRATION NUMBER: 34,637
REFERENCE/DOCKET NUMBER: 2054/22
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-248-7477
TELEFAX: 617-248-7100
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 250 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-461-838-2

Query Match 41.9%; Score 981.5; DB 1; Length 250;
Best Local Similarity 76.8%; Pred. No. 1.6e-73;
Matches 185; Conservative 22; Mismatches 33; Indels 1; Gaps 1;

QY 1 QIOLVSGPELKKPGETVYKISCKASGYPTNYGMNWKAPGQGLKMGMINITSGESTF 60
Db 3 EIQLVSGPELKKPGETVYKISCKASGYPTNYGMNWKAPGQGLKMGMINITSGESTF 62
QY 61 ADFPKRFPDSLETSANTAYLQINNLKSEDMATYFCARMEVYHGVYPYGGQGLTVVSSG 120
Db 63 AEFKGRFAFSLSTASNTAYLQINNLKSEDMATYFCARMEVYHGVYPYGGQGLTVVSSA- 121
QY 121 GGGSGGGSGGGGSDIQLTQSHKFLSTSVGDRVSIITCKASQDVYNAVAYOQKPGQSPKL 180
Db 122 SSSSGSSSSGSSSDIWMQSPKFMSTSVGDRVSIITCKASQDVYNAVAYOQKPGQSPKL 181
QY 181 LIYSASRYTGVPSRFTGSGSPDFTTISVQAEIDLAVYFCQGHFRPTFTGSGTKLEI 240
Db 182 LIWTSRHTGVPSRFTGSGSDTYTLTISVQAEIDLAVYFCQGHFRPTFTGSGTKLEI 241
QY 241 K 241
Db 242 K 242

RESULT 15
US-08-461-386-2
Sequence 2, Application US/08461386
Patent No. 5837846
GENERAL INFORMATION:
APPLICANT: Huston, James S.
APPLICANT: Oppermann, Hermann
APPLICANT: Houston, L. L.
APPLICANT: Ring, David B.
TITLE OF INVENTION: Biosynthetic Binding Proteins For
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:

ADDRESSEE: Testa, Hurwitz & Thibault/Patent Department
STREET: Exchange Place, 53 State Street
CITY: Boston
STATE: Massachusetts
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/461,386
FILING DATE:
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: Kelley, Robin D.
REGISTRATION NUMBER: 34,637
REFERENCE/DOCKET NUMBER: 2054/22
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-248-7477
TELEFAX: 617-248-7100
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 250 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-461-386-2

Query Match 41.9%; Score 981.5; DB 2; Length 250;
Best Local Similarity 76.8%; Pred. No. 1.6e-73;
Matches 185; Conservative 22; Mismatches 33; Indels 1; Gaps 1;

QY 1 QIOLVSGPELKKPGETVYKISCKASGYPTNYGMNWKAPGQGLKMGMINITSGESTF 60
Db 3 EIQLVSGPELKKPGETVYKISCKASGYPTNYGMNWKAPGQGLKMGMINITSGESTF 62
QY 61 ADFPKRFPDSLETSANTAYLQINNLKSEDMATYFCARMEVYHGVYPYGGQGLTVVSSG 120
Db 63 AEFKGRFAFSLSTASNTAYLQINNLKSEDMATYFCARMEVYHGVYPYGGQGLTVVSSA- 121
QY 121 GGGSGGGSGGGGSDIQLTQSHKFLSTSVGDRVSIITCKASQDVYNAVAYOQKPGQSPKL 180
Db 122 SSSSGSSSSGSSSDIWMQSPKFMSTSVGDRVSIITCKASQDVYNAVAYOQKPGQSPKL 181
QY 181 LIYSASRYTGVPSRFTGSGSPDFTTISVQAEIDLAVYFCQGHFRPTFTGSGTKLEI 240
Db 182 LIWTSRHTGVPSRFTGSGSDTYTLTISVQAEIDLAVYFCQGHFRPTFTGSGTKLEI 241
QY 241 K 241
Db 242 K 242

Search completed: April 20, 2005, 07:08:36
Job time : 21.0918 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model1

Run on: April 20, 2005, 07:02:38 ; Search time 56.8451 Seconds
(without alignments)
2590.199 Million cell updates/sec

Title: US-09-596-774-7
Perfect score: 2345
Sequence: 1 QIOLVSGPELKKPGETVKI.....LSTATKDYALHMQLAPR 443

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1421835 seqs, 332370683 residues

Total number of hits satisfying chosen parameters: 1421835

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications NA:
1: /cgn2_6/ptodata/1/pubppaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubppaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubppaa/US06_PUBCOMB.pep:*
4: /cgn2_6/ptodata/1/pubppaa/US07_NEW_PUB.pep:*
5: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
6: /cgn2_6/ptodata/1/pubppaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubppaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubppaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubppaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubppaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubppaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubppaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubppaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubppaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubppaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubppaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/1/pubppaa/US10_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubppaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/1/pubppaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/1/pubppaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	1249	53.3	631	US-10-120-198B-2
2	1015.5	43.3	443	US-10-006-771A-2
3	1015.5	43.3	443	US-10-006-771A-2
4	961.5	41.9	250	US-09-887-853-2
5	969.5	41.3	250	US-10-683-547-2
6	930	39.7	444	US-08-812-393A-2
7	928	39.6	449	US-09-774-681-2
8	917.5	39.1	505	US-10-239-656-79
9	913	38.9	267	US-09-766-543-10
10	903.5	38.5	240	US-10-127-890-148
11	903.5	38.5	240	US-10-127-890-148
12	897	38.3	276	US-09-766-543-12
13	889.5	37.9	541	US-10-071-485-85

14	888.5	37.9	267	US-10-071-485-2	Sequence 2, Appl1
15	888.5	37.9	711	US-10-071-485-90	Sequence 90, Appl1
16	886.5	37.8	503	US-10-239-656-75	Sequence 75, Appl1
17	875.5	37.3	331	US-10-059-261-169	Sequence 169, Appl1
18	872.5	37.2	243	US-10-879-994-10	Sequence 10, Appl1
19	872.5	37.2	243	US-10-879-994-10	Sequence 10, Appl1
20	859	36.6	622	US-10-452-10	Sequence 2, Appl1
21	857	36.5	248	US-10-378-832A-2	Sequence 1104, Ap
22	857	36.5	248	US-09-880-748-1104	Sequence 1104, Ap
23	854.5	36.4	251	US-10-293-418-1921	Sequence 1921, Ap
24	854.5	36.4	251	US-09-880-748-1921	Sequence 1921, Ap
25	854.5	36.4	503	US-10-239-656-77	Sequence 77, Appl1
26	851	36.3	240	US-10-096-246-35	Sequence 35, Appl1
27	851	36.3	240	US-10-096-246-35	Sequence 35, Appl1
28	849.5	36.2	240	US-10-239-656-73	Sequence 73, Appl1
29	848	36.2	240	US-10-071-485-91	Sequence 91, Appl1
30	848	36.2	250	US-09-880-748-932	Sequence 932, App
31	846.5	36.1	249	US-10-293-418-932	Sequence 932, App
32	846.5	36.1	249	US-09-880-748-926	Sequence 926, App
33	846.5	36.1	249	US-10-293-418-926	Sequence 926, App
34	846	36.1	248	US-09-880-748-1446	Sequence 1446, Ap
35	846	36.1	248	US-10-293-418-1446	Sequence 1446, Ap
36	845	36.0	248	US-09-880-748-1008	Sequence 1008, Ap
37	845	36.0	248	US-09-880-748-1008	Sequence 1008, Ap
38	845	36.0	248	US-10-293-418-1008	Sequence 1008, Ap
39	845	36.0	248	US-09-880-748-918	Sequence 918, App
40	842.5	35.9	249	US-10-293-418-918	Sequence 918, App
41	842.5	35.9	249	US-09-984-186-18	Sequence 18, Appl1
42	841.5	35.9	249	US-10-237-667-18	Sequence 18, Appl1
43	841.5	35.9	249	US-10-237-667-18	Sequence 18, Appl1
44	841.5	35.9	249	US-10-237-667-18	Sequence 18, Appl1
45	841.5	35.9	249	US-10-237-667-18	Sequence 18, Appl1

ALIGNMENTS

RESULT 1
US-10-120-198B-2
Sequence 2, Application US/10120198B
Publication No. US20030215427A1
GENERAL INFORMATION:
APPLICANT: Janssen, Michael
TITLE OF INVENTION: CE7-SPECIFIC REDIRECTED IMMUNE CELLS
FILE REFERENCE: 1954-337
CURRENT APPLICATION NUMBER: US/10/120,198B
CURRENT FILING DATE: 2002-04-11
PRIOR APPLICATION NUMBER: 60/282,859
PRIOR FILING DATE: 2001-04-11
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn version 3.1
SEQ ID NO 2
LENGTH: 631
TYPE: PRT
ORGANISM: artificial sequence
FEATURE:
OTHER INFORMATION: mouse-human chimera
US-10-120-198B-2

Query Match 53.3%; Score 1249; DB 15; Length 631;
Best Local Similarity 44.1%; Pred. No. 1,3e-79;
Matches 274; Conservative 60; Mismatches 96; Indels 192; Gaps 13;
QY 1 QIOLVSGPELKKPGETVKISCAASGYPFTNNGNNVYKQAPGGLKMKMGWINSTGASTF 60
23 QVLOQPGALVVRPGASVSKSCASGYTFGTGMMHWKQRPGLGLEWIGERINSNGRTNY 82
QY 61 ADPFGKGFDSLSNTAYLQINNLKSEDMATYFCARMEVYHG---YVPMGCGTTVTY 117
DB 83 NERKSKATLTVDSSSTTAPMQLSGLSBSAYFCAR--DYTGTSYNPDYWGGLTLTV 140
QY 118 SSGGSGGSGGSGGSGGSDIQLTQSHKPLSTSVGDRVSTTCASQDVYNAVAYQOKPGOS 177

```

Db      141  SSGGGSGGGGGGGSDIOMTQSSSFSVSLGDRVITITCKANEDINNRLAMYOQTGNS 200
Qy      178  PKLLIYASAKRYTGPBSRFTGSGSPDFTTITSSVQADLAVYFCQGFRTPTFTGSGTK 237
Db      201  PRLISGATNLVTVGSPSRFSGSGSKDYLTITSLQADDFATYVCOQYMSPTFTGSGTE 260
Qy      238  LEIK-----ALEISNVWFSSVPLVQ-- 260
Db      261  LEIKVEKPSDKTHCPCPAPBELLGSPVFLFPKPDITMISTKTPVTCVVDVSHED 320
Qy      261  -----KVNSTTKP-----VLR-----TPS 275
Db      321  PEVKNMVDVGVENAKTKPREQYNSTYRVSVLTVLHODWLNKGEYCKSVSKALPA 380
Qy      276  PVHPT---GTSOPQRP-----DCRPGS 296
Db      381  PIKRTISKAKQPPRPVYTLPPSRDELTKQVSLTCLVKGFPSDIAVEMESNGQPENN 440
Qy      297  VKGTG--LDPLEDPKLCYLDLIFY----- 321
Db      441  YKTPPVLD-----SDGSFELYSKLTVDSRMQGNVFGSVNHEALNHYTQKS 490
Qy      322  -----GVII--TALYLAKFSRSAAETANLDPNQLYNELNIGRRE 361
Db      491  LSLSPKMALIVLGVAIGLLIFIGLIFFRVYFSRSADAPAYQOGQOLYNELNIGRRE 550
Qy      362  YDVEKKRARPENGGKQRRRNPOEGVYNALOKDMAEYSEIGTGERRGKHGDL 421
Db      551  YDVLDKRRGRPEWGGK--PRKNPOEGLYNELQDKMAEYSEIGMGERRGKHGDL 609
Qy      422  QGLSTATDYDALHMOTLAPR 443
Db      610  QGLSTATDYDALHMOTLAPR 631

RESULT 2
US-10-006-773-2
; Sequence 2, Application US/10006773
; Publication No. US20020132983A1
; GENERAL INFORMATION:
; APPLICANT: Jungmans, Richard P.
; TITLE OF INVENTION: Antibodies as Chimeric Effector Cell Receptors Against Tumor Anti-
; FILE REFERENCE: 003
; CURRENT APPLICATION NUMBER: US/10/006, 773
; CURRENT FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: 60/250, 089
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 443
; TYPE: PRT
; ORGANISM: Homo sapiens and Mus sp.
US-10-006-773-2

```

```

Query Match      43.3%; Score 1015.5; DB 13; Length 443;
Best Local Similarity 52.0%; Pred. No. 2,4e-63;
Matches 225; Conservative 25; Mismatches 58; Indels 125; Gaps 8;

```

```

Qy      131  GGGSDIQLTQSHKFLTSVGDVRSITCKASQDVYNAVAMVYQKFGSPKLLIYSASRYT 190
Db      16   GVHSDIQLTQSPSLASVGDVRYTITCKASQDVYNAVAMVYQKFGKAPKLLIYWTSTRT 75
Qy      191  GVPSRFTGSGSPDFTTITSSVQADLAVYFCQGH--FRTPTFTGSGTKLEIK----- 241
Db      76   GVPSRFTGSGSDPFTTITSLQPEDATATYVCOQYSLYR--SFGQGTVEIKRGGSGSG 132
Qy      242  -----ALEISNV----- 249
Db      133  GSGSGSGSEVOLVESGGVAVQPSRLRLSCSASGDFTTYMWSWVRQAPGKLEWIGE 192
Qy      250  -----MYFSS 254

```

```

Db      193  HPDSSTINAPSLKDRFTTISRDNAKNTLFLQMSLRPEDTGYVFCASLYFGFPWFAYWG 252
Qy      255  VVPLQVKNSTTKKVLRTSPVHPGTSP--QRPEDCRP--GSYKGTGLDPLEDPKL 310
Db      253  GTPVTVSAAKPTTTPARPPTPA-PTIASQPLSLRPEARPAAGAVHTRGDLPALDPKL 311
Qy      311  CYLDDGLIFYGVIITLALYLRKFSRSAAETANLQPNQLYNELNIGRREYVLEKRA 370
Db      312  CYLDDGLIFYGVIITLALFLVKFSRSABPPAYQOGQOLYNELNIGRREYVLDKRRG 371
Qy      371  RDEPMGKQRRRNPOEGVYNALOKDMAEYSEIGTGERRGKHGDLGYOGSLSTATKD 430
Db      372  RDEPMGK--PRKNPOGLVNELQDKMAEYSEIGMGERRGKHGDLGYOGSLSTATKD 430
Qy      431  TYDALHMOTLAPR 443
Db      431  TYDALHMOTLAPR 443

```

```

RESULT 3
US-10-006-771A-2
; Sequence 2, Application US/10006771A
; Publication No. US20020165360A1
; GENERAL INFORMATION:
; APPLICANT: Jungmans, Richard P.
; TITLE OF INVENTION: Chimeric Effector Cell Receptors Against Carcinoembryonic Antigen
; FILE REFERENCE: 002
; CURRENT APPLICATION NUMBER: US/10/006, 771A
; CURRENT FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: 60/250, 090
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 443
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-006-771A-2

```

```

Query Match      43.3%; Score 1015.5; DB 13; Length 443;
Best Local Similarity 52.0%; Pred. No. 2,4e-63;
Matches 225; Conservative 25; Mismatches 58; Indels 125; Gaps 8;

```

```

Qy      131  GGGSDIQLTQSHKFLTSVGDVRSITCKASQDVYNAVAMVYQKFGSPKLLIYSASRYT 190
Db      16   GVHSDIQLTQSPSLASVGDVRYTITCKASQDVYNAVAMVYQKFGKAPKLLIYWTSTRT 75
Qy      191  GVPSRFTGSGSPDFTTITSSVQADLAVYFCQGH--FRTPTFTGSGTKLEIK----- 241
Db      76   GVPSRFTGSGSDPFTTITSLQPEDATATYVCOQYSLYR--SFGQGTVEIKRGGSGSG 132
Qy      242  -----ALEISNV----- 249
Db      133  GSGSGSGSEVOLVESGGVAVQPSRLRLSCSASGDFTTYMWSWVRQAPGKLEWIGE 192
Qy      250  -----MYFSS 254
Db      193  HPDSSTINAPSLKDRFTTISRDNAKNTLFLQMSLRPEDTGYVFCASLYFGFPWFAYWG 252
Qy      255  VVPLQVKNSTTKKVLRTSPVHPGTSP--QRPEDCRP--GSYKGTGLDPLEDPKL 310
Db      253  GTPVTVSAAKPTTTPARPPTPA-PTIASQPLSLRPEARPAAGAVHTRGDLPALDPKL 311
Qy      311  CYLDDGLIFYGVIITLALYLRKFSRSAAETANLQPNQLYNELNIGRREYVLEKRA 370
Db      312  CYLDDGLIFYGVIITLALFLVKFSRSABPPAYQOGQOLYNELNIGRREYVLDKRRG 371
Qy      371  RDEPMGKQRRRNPOEGVYNALOKDMAEYSEIGTGERRGKHGDLGYOGSLSTATKD 430
Db      372  RDEPMGK--PRKNPOEGLYNELQDKMAEYSEIGMGERRGKHGDLGYOGSLSTATKD 430
Qy      431  TYDALHMOTLAPR 443

```


MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/812.393A
FILING DATE: 05-MAR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Murashige, Kate H
REGISTRATION NUMBER: 29,959
REFERENCE/DOCKET NUMBER: 31333-20001.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-887-1500
TELEFAX: 202-822-0168
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 444 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
US-08-812-393A-2

Query Match 39.7%; Score 930; DB 8; Length 444;
Best Local Similarity 48.2%; Pred. No. 2.6e-57;
Matches 217; Conservative 50; Mismatches 145; Indels 38; Gaps 14;

QY 3 QLVOSGPE--LKKRGETVTKISCKASGVPTNYGMNVKQAPGGLKMMGMINTSGSTF 60
DB 24 QVQVQSPLSLVLOEGBNAELQCSFS--IFTNQ-VQMFYQRPGRVLSLNPSTGKOS-- 78
QY 61 ADPEKGFDPFSLTSANTAVLQINNLSKSEDMATVFCARMEVYHGVYVPGQGTVTYSSG 120
DB 79 -----GRLTSTTVTKERRSSLHSSQITDSCTYLCA NSGCSNAKLTFGKGLTSVKSG 133
QY 121 GGGSGGGSGGGSGSDIQLTQSHKFLSTVGDRVSTTCASQDYNAVANAYQKFGQSPKL 180
DB 134 GGGSGGGSGGGSGSEAAVTQSPRNKVAVTGKVTLSQNTNN--HNNMYWYKODTGHGLRL 192
QY 181 LIYS---ASSRYTVGVSFRFGS--GSGPDFFTTISVQAEIDLAVYFC---QQHRTPTFTFG 233
DB 193 IHSYAGSTEKEDIIDGYKASRPSQENFSLILELATPSCTSVYFCASGETGTNERLFPFG 252
QY 234 SGTLEIKALEISNSVMYFSSVVPVLQKVNSTTKPVLRTPSPVHPTGTSQORPEDCRP 293
DB 253 HGTKLSTVLT---SNSIMYFHFVFLPAKPTTT--PAPRPPTPA--PTIASQPL---SLRP 304
QY 294 RGSVKGTDLPLEDPKLCYLLDGLFTYGVITLALYLRKAFSRSATANLQDPNQLYNE 353
DB 305 SSS-----RDPKLCYLLDGLFTYGVITLALYLRKAFSRSADAPAYVQOGQNLVNE 355
QY 354 INLGRREYDVLLEKRRADPEMGKQOQRRNPQEGVYNALQDKMAEAYSEIGTKGERRR 413
DB 356 INLGRREYDVLQKRRGRPEMGK--PRKKNPQEGLYNELQDKMAEAYSEIGTKGERRR 414
QY 414 GKGHDLGYOGLSTATKDTYDALHMQLAPR 443
DB 415 GKGHDLGYOGLSTATKDTYDALHMQLAPR 444

RESULT 7
US-09-774-681-2
Sequence 2, Application US/09774681
Publication No. US20030208780A1
GENERAL INFORMATION:
APPLICANT: Sunol Molecular Corporation
APPLICANT: Sherman, Linda

APPLICANT: Lustgarten, Joseph
TITLE OF INVENTION: RECOMBINANT CONSTRUCTS ENCODING T CELL
TITLE OF INVENTION: RECEPTORS SPECIFIC FOR HUMAN HLA-RESTRICTED TUMOR ANTIGENS
FILE REFERENCE: 31333-20001.01
CURRENT APPLICATION NUMBER: US/09/774,681
CURRENT FILING DATE: 2001-02-01
PRIOR APPLICATION NUMBER: US 08/812,393
PRIOR FILING DATE: 1997-03-05
PRIOR APPLICATION NUMBER: US 60/012,845
PRIOR FILING DATE: 1996-03-05
NUMBER OF SEQ ID NOS: 65
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 449
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Deduced amino acid derivative of effective T cell
US-09-774-681-2

Query Match 39.6%; Score 928; DB 10; Length 449;
Best Local Similarity 48.2%; Pred. No. 3.6e-57;
Matches 217; Conservative 50; Mismatches 145; Indels 38; Gaps 14;

QY 3 QLVOSGPE--LKKRGETVTKISCKASGVPTNYGMNVKQAPGGLKMMGMINTSGSTF 60
DB 24 QVQVQSPLSLVLOEGBNAELQCSFS--IFTNQ-VQMFYQRPGRVLSLNPSTGKOS-- 78
QY 61 ADPEKGFDPFSLTSANTAVLQINNLSKSEDMATVFCARMEVYHGVYVPGQGTVTYSSG 120
DB 79 -----GRLTSTTVTKERRSSLHSSQITDSCTYLCA NSGCSNAKLTFGKGLTSVKSG 133
QY 121 GGGSGGGSGGGSGSDIQLTQSHKFLSTVGDRVSTTCASQDYNAVANAYQKFGQSPKL 180
DB 134 GGGSGGGSGGGSGSEAAVTQSPRNKVAVTGKVTLSQNTNN--HNNMYWYKODTGHGLRL 192
QY 181 LIYS---ASSRYTVGVSFRFGS--GSGPDFFTTISVQAEIDLAVYFC---QQHRTPTFTFG 233
DB 193 IHSYAGSTEKEDIIDGYKASRPSQENFSLILELATPSCTSVYFCASGETGTNERLFPFG 252
QY 234 SGTLEIKALEISNSVMYFSSVVPVLQKVNSTTKPVLRTPSPVHPTGTSQORPEDCRP 293
DB 253 HGTKLSTVLT---SNSIMYFHFVFLPAKPTTT--PAPRPPTPA--PTIASQPL---SLRP 304
QY 294 RGSVKGTDLPLEDPKLCYLLDGLFTYGVITLALYLRKAFSRSATANLQDPNQLYNE 353
DB 305 SSS-----RDPKLCYLLDGLFTYGVITLALYLRKAFSRSADAPAYVQOGQNLVNE 355
QY 354 INLGRREYDVLLEKRRADPEMGKQOQRRNPQEGVYNALQDKMAEAYSEIGTKGERRR 413
DB 356 INLGRREYDVLQKRRGRPEMGK--PRKKNPQEGLYNELQDKMAEAYSEIGTKGERRR 414
QY 414 GKGHDLGYOGLSTATKDTYDALHMQLAPR 443
DB 415 GKGHDLGYOGLSTATKDTYDALHMQLAPR 444

RESULT 8
US-10-239-656-79
Sequence 79, Application US/10239656
Publication No. US2004003839A1
GENERAL INFORMATION:
APPLICANT: KUPER, PETER
APPLICANT: RIETMULLER, GERT
APPLICANT: LUTTERBOUSE, RALF
APPLICANT: BORSCHERT, KATRIN
APPLICANT: KISCHEL, ROMAN
APPLICANT: MAYER, MONIKA
APPLICANT: HOMERISTER, ROBERT
TITLE OF INVENTION: MULTIFUNCTIONAL POLYPEPTIDES COMPRISING A BINDING SITE
TITLE OF INVENTION: TO AN EPITOPE OF THE NKGD2 RECEPTOR COMPLEX
FILE REFERENCE: 029976/0106

OTHER INFORMATION: IL-2-antibody fusions
US-09-766-543-10
Query Match 38.9%; Score 913; DB 9; Length 267;
Best Local Similarity 68.4%; Pred. No. 2.3e-56;
Matches 167; Conservative 37; Mismatches 36; Indels 4; Gaps 2;

Query Match 39.1%; Score 917.5; DB 15; Length 505;
Best Local Similarity 50.5%; Pred. No. 2.2e-56;
Matches 185; Conservative 25; Mismatches 31; Indels 125; Gaps 4;

OTHER INFORMATION: Description of Artificial Sequence: Synthetic 3810xps-
US-10-239-656-79

OTHER INFORMATION: 23 bp specific single chain Fv

SEQUENCE: 101VOSGPELKKGGTETVTKISCKASGYPTNYGMNWKQAPGQGLKMMGINTSTGESEF 60
128 QVQLQSGPELKKGGTETVTKISCKASGYPTNYGMNWKQAPGQGLKMMGINTSTGESEF 187
61 ADPKGRPFSLKTSANTAYLQINNLSKEDMATYFCAR----- 98
188 GDDPKGRPFSLKTSANTAYLQINNLSKEDMATYFCAR----- 247
99 -----W----- 99

SEQUENCE: 100-----EVH-----GYPTNGGCTTTVSSGG 121
248 EVQLLESGGGLVPGQSLKLSCKASGPDPSRYSWVQAPGKLEIGELINDPSSTNY 307
100 -----EVH-----GYPTNGGCTTTVSSGG 121
308 TPSSKDRPFISRDNAKNTLYQMSKVRSEDTALYCARLQGMGQYFDYGGCTTTVSSGG 367
122 GSGSGSGSGSGSDIQLTQSHKFLSTSVGRVSTICKASQDVYNA-----VAMVQKPKG 175
368 GSGSGSGSGSGSDIQLTQSHKFLSTSVGRVSTICKASQDVYNA-----VAMVQKPKG 427
176 QSPRLIYASASRYTVGVPFRFTGSGSDPFTTISVQAEADLAIVFCQOHFRPTFTGSG 235
428 QPRLIYMASTRSGVDPFRFTGSGSDPFTTISVQAEADLAIVFCQOHFRPTFTGSG 487

SEQUENCE: 236 TKLEIK 241
488 TKLEIK 493

RESULT 9
US-09-766-543-10
Sequence 10, Application US/09766543
Patent No. US20020041865A1
GENERAL INFORMATION:
APPLICANT: AuscIn, Richard
APPLICANT: Kwok, Cheuk S.
TITLE OF INVENTION: METHODS FOR TREATING TUMORS
FILE REFERENCE: PP01679.002
CURRENT APPLICATION NUMBER: US/09/766,543
CURRENT FILING DATE: 2000-01-19
PRIOR APPLICATION NUMBER: 60/177,258
PRIOR FILING DATE: 2000-01-20
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 10
LENGTH: 267
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: 520C9
OTHER INFORMATION: humanized single-chain antibody used in the

OTHER INFORMATION: IL-2-antibody fusions
US-09-766-543-10
Query Match 38.9%; Score 913; DB 9; Length 267;
Best Local Similarity 68.4%; Pred. No. 2.3e-56;
Matches 167; Conservative 37; Mismatches 36; Indels 4; Gaps 2;

Query Match 39.1%; Score 917.5; DB 15; Length 505;
Best Local Similarity 50.5%; Pred. No. 2.2e-56;
Matches 185; Conservative 25; Mismatches 31; Indels 125; Gaps 4;

OTHER INFORMATION: Description of Artificial Sequence: Synthetic 3810xps-
US-10-239-656-79

OTHER INFORMATION: 23 bp specific single chain Fv

SEQUENCE: 101VOSGPELKKGGTETVTKISCKASGYPTNYGMNWKQAPGQGLKMMGINTSTGESEF 60
128 QVQLQSGPELKKGGTETVTKISCKASGYPTNYGMNWKQAPGQGLKMMGINTSTGESEF 187
61 ADPKGRPFSLKTSANTAYLQINNLSKEDMATYFCAR----- 98
188 GDDPKGRPFSLKTSANTAYLQINNLSKEDMATYFCAR----- 247
99 -----W----- 99

SEQUENCE: 100-----EVH-----GYPTNGGCTTTVSSGG 121
248 EVQLLESGGGLVPGQSLKLSCKASGPDPSRYSWVQAPGKLEIGELINDPSSTNY 307
100 -----EVH-----GYPTNGGCTTTVSSGG 121
308 TPSSKDRPFISRDNAKNTLYQMSKVRSEDTALYCARLQGMGQYFDYGGCTTTVSSGG 367
122 GSGSGSGSGSGSDIQLTQSHKFLSTSVGRVSTICKASQDVYNA-----VAMVQKPKG 175
368 GSGSGSGSGSGSDIQLTQSHKFLSTSVGRVSTICKASQDVYNA-----VAMVQKPKG 427
176 QSPRLIYASASRYTVGVPFRFTGSGSDPFTTISVQAEADLAIVFCQOHFRPTFTGSG 235
428 QPRLIYMASTRSGVDPFRFTGSGSDPFTTISVQAEADLAIVFCQOHFRPTFTGSG 487

SEQUENCE: 236 TKLEIK 241
488 TKLEIK 493

RESULT 10
US-10-127-890-148
Sequence 148, Application US/10127890
Publication No. US20030166196A1
GENERAL INFORMATION:
APPLICANT: Better, Marc D.
Carroll, Stephen F.
Studnick, Gary M.
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
Proteins
NUMBER OF SEQUENCES: 173
CORRESPONDENCE ADDRESS:
ADDRESSEE: McAndrew, Held & Malloy, Ltd.
STREET: 500 West Madison Street, 34th Floor
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60661
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/127,890
FILING DATE: 23-Apr-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/646,360
FILING DATE: 13-MAY-1996
APPLICATION NUMBER: PCT/US94/05348
FILING DATE: 12-MAY-1994
APPLICATION NUMBER: US 08/064,691
FILING DATE: 12-MAY-1993
APPLICATION NUMBER: US 07/988,430
FILING DATE: 09-DEC-1992
APPLICATION NUMBER: US 07/901,707
FILING DATE: 19-JUN-1992
APPLICATION NUMBER: US 07/787,567
FILING DATE: 04-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: McNicholas, Janet M.
REGISTRATION NUMBER: 32,918
REFERENCE/DOCKET NUMBER: 200-70.P4
TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/707-8889
TELEFAX: 312/707-9155
TELEX: 650 388-1248
INFORMATION FOR SEQ ID NO: 148:
SEQUENCE CHARACTERISTICS:
LENGTH: 240 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 148:
US-10-127-880-148

Query Match 38.5%; Score 903.5; DB 14; Length 240;
Best Local Similarity 69.3%; Pred. No. 9,4e-56;
Matches 167; Conservative 32; Mismatches 41; Indels 1; Gaps 1;

QY 1 QIOLVSGPELKKRGETYKISCKASGYPTNYGMNVKQAPGGGLKMMGMINTSGESTF 60
DB 1 EIQLVSGGGLVPRGSGVRSICASGYPTNYGMNVKQAPGGGLKMMGMINTHTGEPT 60
QY 61 ADFFKGRFDFSLTSSANTAYLQINNLKSEDMATYFCARWEVYHGVPYWGQGTIVVSSG 120
DB 61 ADFFKGRFDFSLDSSKNTAYLQINSLRAEDTAVYFCR-RGYDWPYDVGQGTIVVSSG 119
QY 121 GGGSGGGSGGGSDIQLTQSHKFLSTVGDVSTICKASQDVYNAVAYQKPGQSPKL 180
DB 120 GGGSGGGSGGGSDIQMTQSPSLASVGDVITTCRASQDINSYLSWFOQKPGKAPKT 179
QY 181 LIYASASRYTVGVRPFGSGGSDPTFTTSSVOAEDLAVFCQGHFRTPTFGSGTKLEI 240
DB 180 LIYANRLBSGVSRFGSGGSDTYTLTSSLYEDFGIYCCQYDESPWTFGGTKLEW 239
QY 241 K 241
DB 240 K 240

RESULT 11
US-10-717-243-148
Sequence 148, Application US/10717243
Publication No. US20050054835A1
GENERAL INFORMATION:
APPLICANT: Better, Marc D.
Carroll, Stephen F.
Studinka, Gary M.
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
Proteins
NUMBER OF SEQUENCES: 169
CORRESPONDENCE ADDRESS:
ADDRESSEE: McAndrews, Held & Malloy, Ltd.
STREET: 500 West Madison Street, 34th floor
CITY: Chicago
STATE: Illinois
COUNTRY: USA
ZIP: 60661

COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10717,243
FILING DATE: 18-Nov-2003
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/839,765
FILING DATE: 15-APR-1997
APPLICATION NUMBER: US 08/425,336
FILING DATE: 18-APR-1995
APPLICATION NUMBER: US 08/064,691
FILING DATE: 12-MAY-1993
APPLICATION NUMBER: US 07/988,430
FILING DATE: 09-DEC-1992

APPLICATION NUMBER: US 07/901,707
FILING DATE: 19-JUN-1992
APPLICATION NUMBER: US 07/787,567
FILING DATE: 04-NOV-1991
ATTORNEY/AGENT INFORMATION:
NAME: McNicholas, Janet M.
REGISTRATION NUMBER: 32,918
REFERENCE/DOCKET NUMBER: 11022US09/200-70.P3.C3
TELECOMMUNICATION INFORMATION:
TELEPHONE: 312/707-8889
TELEFAX: 312/707-9155
TELEX: 650 388-1248
INFORMATION FOR SEQ ID NO: 148:
SEQUENCE CHARACTERISTICS:
LENGTH: 240 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 148:
US-10-717-243-148

Query Match 38.5%; Score 903.5; DB 17; Length 240;
Best Local Similarity 69.3%; Pred. No. 9,4e-56;
Matches 167; Conservative 32; Mismatches 41; Indels 1; Gaps 1;

QY 1 QIOLVSGPELKKRGETYKISCKASGYPTNYGMNVKQAPGGGLKMMGMINTSGESTF 60
DB 1 EIQLVSGGGLVPRGSGVRSICASGYPTNYGMNVKQAPGGGLKMMGMINTHTGEPT 60
QY 61 ADFFKGRFDFSLTSSANTAYLQINNLKSEDMATYFCARWEVYHGVPYWGQGTIVVSSG 120
DB 61 ADFFKGRFDFSLDSSKNTAYLQINSLRAEDTAVYFCR-RGYDWPYDVGQGTIVVSSG 119
QY 121 GGGSGGGSGGGSDIQLTQSHKFLSTVGDVSTICKASQDVYNAVAYQKPGQSPKL 180
DB 120 GGGSGGGSGGGSDIQMTQSPSLASVGDVITTCRASQDINSYLSWFOQKPGKAPKT 179
QY 181 LIYASASRYTVGVRPFGSGGSDPTFTTSSVOAEDLAVFCQGHFRTPTFGSGTKLEI 240
DB 180 LIYANRLBSGVSRFGSGGSDTYTLTSSLYEDFGIYCCQYDESPWTFGGTKLEW 239
QY 241 K 241
DB 240 K 240

RESULT 12
US-09-766-543-12
Sequence 12, Application US/09766543
Patent No. US20020041865A1
GENERAL INFORMATION:
APPLICANT: Austin, Richard
APPLICANT: Kwok, Cheuk S.
APPLICANT: Ring, David B.
TITLE OF INVENTION: METHODS FOR TREATING TUMORS
FILE REFERENCE: P01679,002
CURRENT APPLICATION NUMBER: US/09/766,543
PRIOR FILING DATE: 2000-01-19
PRIOR APPLICATION NUMBER: 60/177,258
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 12
LENGTH: 276
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: H520C9sfv plus
US-09-766-543-12

Query Match 38.3%; Score 897; DB 9; Length 276;
Best Local Similarity 69.7%; Pred. No. 3.1e-55;

Matches 168; Conservative 34; Mismatches 35; Indels 4; Gaps 2;

Qy 1 QIOLVOSGPELKKPGETVTKISCKASGYPTNYGNMVKQAPGQGLKMMGWINTSGESTF 60
Db 23 EIQLVOSGPEVKKKQKASVKISCKASGYTFANYNMVKQAPGQGLKMMGWINTYGSSTY 82

Qy 61 ADPFKGRFDSLETSANTAYIQINNLSKSEDMATTFCAWMEVYHGVYVWGQTTVTVSSG 120
Db 83 ADPFKERTFLDSTSTAHLEISLSREDDATYFCAR--RFGFA-YWGQGLTVTVSSG 138

Qy 121 GGGSGGGSGGGSDIQLTQSHKFLSTSVGDRVSTTCASQDVYNAVAYQKPGQSPKL 180
Db 139 GGGSGGGSGGGSDIQLTQSPSSLSASVGDVRLTTCRASODIGNSLTWLOQKPKTKR 198

Qy 181 LIYASASRYTGVPSPRFTGSGSGPDFTFTISSVOAEDLAVYFCQOHFRTPFTFGSGTKLEI 240
Db 199 LIYATSSLDGVPSPRFTGSGSGPDFTFTISSVOAEDLAVYFCQOHFRTPFTFGSGTKLEI 258

Qy 241 K 241
Db 259 K 259

RESULT 13
US-10-071-485-85
; Sequence 85, Application US/10071485
; Publication No. US20030099648A1
; GENERAL INFORMATION:
; APPLICANT: Buyee, Marie-Ange
; APPLICANT: Sablon, Erwin
; TITLE OF INVENTION: INTERFERON-gamma-BINDING MOLECULES FOR TREATING SEPTIC
; TITLE OF INVENTION: SHOCK,
; TITLE OF INVENTION: CACHEXIA, IMMUNE DISEASES AND SKIN DISORDERS
; FILE REFERENCE: INNS:015
; CURRENT APPLICATION NUMBER: US/10/071.485
; PRIOR FILING DATE: 2002-02-07
; PRIOR APPLICATION NUMBER: 09/485,737
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: PCT/EP 98/05165
; PRIOR FILING DATE: 1998-08-14
; PRIOR APPLICATION NUMBER: EPO 98870139.7
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: EPO 97870122.5
; PRIOR FILING DATE: 1997-08-18
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 85
; LENGTH: 541
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: SYNTHETIC
US-10-071-485-85

Query Match 37.9%; Score 889.5; DB 14; Length 541;
Best Local Similarity 67.7%; Pred. No. 2.3e-54;
Matches 168; Conservative 30; Mismatches 47; Indels 3; Gaps 2;

Qy 1 QIOLVOSGPELKKPGETVTKISCKASGYPTNYGNMVKQAPGQGLKMMGWINTSGESTF 60
Db 1 QVQLVOSGSELKKRKGASVKISCKASGYTFDYGMMVKQAPGQGLKMMGWINTYGSSTY 60

Qy 61 ADPFKGRFDSLETSANTAYIQINNLSKSEDMATTFCAWMEVYHGVYVWGQTTVTVSSG 120
Db 61 VDDPKGRFVPSLDTSVSAAYLQISLSLKAEDATYFCARGFY--AMDYWGQTTVTVSSG 118

Qy 121 GGGSGGGSGGGSDIQLTQSHKFLSTSVGDRVSTTCASQDVYNAVAYQKPGQSPKL 180
Db 119 GGGSGGGSGGGSDIQLTQSPATMSASPGERVTLTCSASSI--SYMFWHQRPQSPRL 177

Qy 181 LIYASASRYTGVPSPRFTGSGSGPDFTFTISSVOAEDLAVYFCQOHFRTPFTFGSGTKLEI 240
Db 178 LIYDTSNLAGSVPARFGSGSGTSLTISLMBEDATYFCQSSSYPTFTFGGQTKLEI 237

Qy 241 KALEISNS 248
Db 238 KRTPLGDT 245

RESULT 14
US-10-071-485-2
; Sequence 2, Application US/10071485
; Publication No. US20030099648A1
; GENERAL INFORMATION:
; APPLICANT: Buyee, Marie-Ange
; APPLICANT: Sablon, Erwin
; TITLE OF INVENTION: INTERFERON-gamma-BINDING MOLECULES FOR TREATING SEPTIC
; TITLE OF INVENTION: SHOCK,
; TITLE OF INVENTION: CACHEXIA, IMMUNE DISEASES AND SKIN DISORDERS
; FILE REFERENCE: INNS:015
; CURRENT APPLICATION NUMBER: US/10/071.485
; PRIOR FILING DATE: 2002-02-07
; PRIOR APPLICATION NUMBER: 09/485,737
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: PCT/EP 98/05165
; PRIOR FILING DATE: 1998-08-14
; PRIOR APPLICATION NUMBER: EPO 98870139.7
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: EPO 97870122.5
; PRIOR FILING DATE: 1997-08-18
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-071-485-2

Query Match 37.9%; Score 888.5; DB 14; Length 267;
Best Local Similarity 69.7%; Pred. No. 1.2e-54;
Matches 168; Conservative 27; Mismatches 43; Indels 3; Gaps 2;

Qy 1 QIOLVOSGPELKKPGETVTKISCKASGYPTNYGNMVKQAPGQGLKMMGWINTSGESTF 60
Db 23 QVQLVOSGSELKKRKGASVKISCKASGYTFDYGMMVKQAPGQGLKMMGWINTYGSSTY 82

Qy 61 ADPFKGRFDSLETSANTAYIQINNLSKSEDMATTFCAWMEVYHGVYVWGQTTVTVSSG 120
Db 83 VDDPKGRFVPSLDTSVSAAYLQISLSLKAEDATYFCARGFY--AMDYWGQTTVTVSSG 140

Qy 121 GGGSGGGSGGGSDIQLTQSHKFLSTSVGDRVSTTCASQDVYNAVAYQKPGQSPKL 180
Db 141 GGGSGGGSGGGSDIQLTQSPATMSASPGERVTLTCSASSI--SYMFWHQRPQSPRL 199

Qy 181 LIYASASRYTGVPSPRFTGSGSGPDFTFTISSVOAEDLAVYFCQOHFRTPFTFGSGTKLEI 240
Db 200 LIYDTSNLAGSVPARFGSGSGTSLTISLMBEDATYFCQSSSYPTFTFGGQTKLEI 259

Qy 241 K 241
Db 260 K 260

RESULT 15
US-10-071-485-90
; Sequence 90, Application US/10071485
; Publication No. US20030099648A1
; GENERAL INFORMATION:
; APPLICANT: Buyee, Marie-Ange
; APPLICANT: Sablon, Erwin
; TITLE OF INVENTION: INTERFERON-gamma-BINDING MOLECULES FOR TREATING SEPTIC
; TITLE OF INVENTION: SHOCK,
; TITLE OF INVENTION: CACHEXIA, IMMUNE DISEASES AND SKIN DISORDERS
; FILE REFERENCE: INNS:015

; CURRENT APPLICATION NUMBER: US/10/071,485
 ; CURRENT FILING DATE: 2002-02-07
 ; PRIOR APPLICATION NUMBER: 09/485,737
 ; PRIOR FILING DATE: 2000-02-14
 ; PRIOR APPLICATION NUMBER: PCT/EP 98/05165
 ; PRIOR FILING DATE: 1998-06-14
 ; PRIOR APPLICATION NUMBER: EPO 98870139.7
 ; PRIOR FILING DATE: 1998-06-18
 ; PRIOR APPLICATION NUMBER: EPO 97870122.5
 ; PRIOR FILING DATE: 1997-08-18
 ; NUMBER OF SEQ ID NOS: 104
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 90
 ; LENGTH: 711
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: SYNTHETIC
 US-10-071-485-90

Query Match 37.9%; Score 888.5; DB 14; Length 711;
 Best Local Similarity 69.7%; Pred. No. 3.6e-54;
 Matches 168; Conservative 27; Mismatches 43; Indels 3; Gaps 2;

QY	1	QIOLVSGPELKKPGETVYKISCKASGYPTNYGMNWKQAPGQGLKMWGINSTGSESTF	60
DB	473	QVQLVSGSSELKKPGASVYKISCKASGYFTDYGMMWVKQAPGQGLKMWGINITYTGSESTY	532
QY	61	ADDFKGRFDFLETSANTAYVQINNLSKSEDMATYFCARWEYHGVPYWGQGTIVTVSSG	120
DB	533	VDDPKGRFVPSLDTSVSAVYLIQISSLKAEADYATYFCARGRFY--AMDYWGQGTIVTVSSG	590
QY	121	GGSGGGGGGGGGSDIQLTOSHKFLSTSVGDVRSITCKASQDVYNAVAVYQKPGQSPKL	180
DB	591	GGSGGGGGGGGGSDIYLTOBPATMSASPERVTLTCSASSI-SYMFVTHQRPGQSPRL	649
QY	181	LIIYASRRYGVPSRFTGSGSGGPDFTFTISSVQAEADLAVYFCQGHFRTPFTFGSGTKLEI	240
DB	650	LIIYDTSMIAGGVPARFSGSGSGTSLTISRMEDEDPATYFCQSSSYPTFGQGTLEI	709
QY	241	K 241	
DB	710	K 710	

Search completed: April 20, 2005, 07:10:39
 Job time : 57.8451 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 20, 2005, 10:01:29 ; Search time 5925.62 Seconds
(without alignments)
3622.520 Million cell updates/sec

Title: US-09-596-774-7

Perfect score: 2345
Sequence: 1 QIOLVQSGPRLKKRGFTVKI.....LSTATKDYALHMQTLAPR 443

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+g2n+model -DEV=xlh
-Q=/cgnt2_1/USPTO_spool_h/US0956774/runatc_20042005_075442_12508/app_query.fasta_1.1230
-DB=GenEmbl -QFMT=faetap -SUFFIX=aprt20.rge -MINMATCH=0.1 -LOOPCL=0 -LOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blousum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=spc -THR MAX=100 -THR MIN=0 -ALIGN=45 -MODE=LOCAL
-OUTFMT=prc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USRR=US0956774.QCGN_1.1_6663@runatc_20042005_075442_12508 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -YGAPOP=10 -YGAPEXT=0.5 -DELop=6 -DELEXT=7

Database :

GenEmbl:.*
1: gb_Da:.*
2: gb_htg:.*
3: gb_in:.*
4: gb_om:.*
5: gb_ov:.*
6: gb_pac:.*
7: gb_ph:.*
8: gb_pl:.*
9: gb_pr:.*
10: gb_ro:.*
11: gb_rts:.*
12: gb_sy:.*
13: gb_un:.*
14: gb_vi:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2345	100.0	1479	6	A47281 Sequence 5
2	1501.5	64.0	1284	6	A93938 Sequence 3
3	1501.5	64.0	3058	6	A93936 Sequence 1
4	1312.5	56.0	2070	6	A57337 Sequence 7

5	1312.5	56.0	2730	6	A57353	A57353 Sequence 23
6	1310.5	55.9	2012	6	A22539	A22539 M.musculus/
7	1310.5	55.9	2012	6	128526	128526 Sequence 13
8	1308	55.8	1862	6	A50995	A50995 Sequence 36
9	1308	55.8	1862	6	AR268367	AR268367 Sequence
10	1308	55.8	1919	6	A50993	A50993 Sequence 34
11	1308	55.8	1919	6	AR268366	AR268366 Sequence
12	1308	55.8	2214	6	A57341	A57341 Sequence 11
13	1308	55.8	2793	6	A57343	A57343 Sequence 11
14	1308	55.8	2793	6	A57351	A57351 Sequence 21
15	1308	55.8	3177	6	A57347	A57347 Sequence 17
16	1305	55.7	1692	6	A50960	A50960 Sequence 1
17	1305	55.7	1692	6	AR268336	AR268336 Sequence
18	1303	55.6	894	12	XXU64991	U64994 Synthetic c
19	1303	55.6	2233	6	A22470	A22470 M.musculus/
20	1303	55.6	2233	6	128521	128521 Sequence 6
21	1302	55.5	794	6	A57333	A57333 Sequence 3
22	1302	55.5	2718	6	A57349	A57349 Sequence 19
23	1302	55.5	2793	6	A57345	A57345 Sequence 15
24	1285.5	54.8	748	6	A22469	A22469 M.musculus
25	1285.5	54.8	748	6	A47277	A47277 Sequence 1
26	1285.5	54.8	748	6	128520	128520 Sequence 4
27	1013	43.2	726	6	AX100174	AX100174 Sequence
28	1011.5	43.1	1422	6	A63768	A63768 Sequence 32
29	1011.5	43.1	1545	6	A63770	A63770 Sequence 34
30	994	42.4	771	6	AX057978	AX057978 Sequence
31	981.5	41.9	909	6	AR007979	AR007979 Sequence
32	981.5	41.9	909	6	AR058994	AR058994 Sequence
33	981.5	41.9	909	6	123444	123444 Sequence 1
34	981.5	41.9	909	6	143364	143364 Sequence 7
35	981.5	41.9	909	6	161419	161419 Sequence 7
36	981.5	41.9	909	6	196033	196033 Sequence 7
37	975.5	41.6	777	6	AX935402	AX935402 Sequence
38	975.5	41.6	987	6	AX935453	AX935453 Sequence
39	962.5	41.0	1956	6	A63774	A63774 Sequence 38
40	962.5	41.0	2079	6	A63776	A63776 Sequence 40
41	960.5	41.0	777	6	AX935404	AX935404 Sequence
42	960.5	41.0	987	6	AX935455	AX935455 Sequence
43	942	40.2	1632	12	AY452127	AY452127 Synthetic
44	942	40.2	1761	12	AY452134	AY452134 Synthetic
45	942	40.2	3195	12	AY452128	AY452128 Synthetic

ALIGNMENTS

RESULT 1	A47281	1479 bp	DNA	linear	PAT 07-MAR-1997
LOCUS	A47281				
DEFINITION	Sequence 5 from Patent WO9530014.				
ACCESSION	A47281				
VERSION	A47281.1 GI:2301300				
KEYWORDS					
SOURCE	unidentified				
ORGANISM	unidentified				
REFERENCE	1 (bases 1 to 1479)				
AUTHORS	Groner, B. and Moritz, D.				
TITLE	BIFUNCTIONAL PROTEIN, PREPARATION AND USE				
JOURNAL	Patent: WO 9530014-A 5 09-NOV-1995;				
COMMENT	CIBA GEIGY AG (CH)				
FEATURES	Other publication ZA 9503440 951102				
source	Location/Qualifiers				
	1..1479				
	/organism="unidentified"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:32644"				
	40..1425				
	/codon_start=1				
	/product="LEADER-SCFV(FRPS):LYT-2 HINGE-ZETA"				
	/protein_id="CAA03050.1"				
	/db_xref="GI:2301301"				
	/translation="MAWVWTLFLMAAAKVPKQIQVQSGPRLKKRGFTVKISCKASG				


```

Oy 1 GlnIleGlnLeuValGlnSerGlyProGlnLeuLysProGlyValThrValLysIle 20
Db 13 CAGGTCCAACTGACAGCTCTGGGGCTGAGCTTGGAAGCTCGGGGCTCACTGAAGCTG 72
Oy 21 SerGlyValAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTrpValLysGlnAla 40
Db 73 TCCTGCAAGGCTTCGGCTACACCTTCACACCTGATGCTGGGTGAAGCAGAGG 132
Oy 41 ProGlyGlnGlyLeuLysTrpMetGlyTyrIleAsnThrSerThrGlyLysSerThrPhe 60
Db 133 CTGGACGAGGCGCTTGAAGTGGAGATTGATCCATATAGTGGTGTACTAAGTAC 192
Oy 61 AlaAspAspPheLysGlyArgPheAspPheSerLeuGlnThrSerAlaAsnThrAlaTyr 80
Db 193 AATGAGAAGTTCAAGACGAGGCGCACACTGACTAGACAAACCTCCAGACAGCCTTAC 252
Oy 81 LeuGlnIleAsnAsnLeuLysSerGlnAspMetAlaThrTyrPheCysAlaArgTrpGlu 100
Db 253 ATGACGCTCAGACGCTGACATCTGAGACTCTGGGGCTATATTATTTGCAAGATACGAT 312
Oy 101 ValTyr--HisGlyTyrValProTyrTrpGlyGlnGlyThrThrValThrValSerSer 119
Db 313 TACTACGGTAGTACTGACTTCTGACTGAGGCGCAAGGACACGCTACCGCTCTCTCCA 372
Oy 120 GlyGlyGlyGlySerGlyValGlySerGlyGlyGlySerAspIleGlnLeuThr 139
Db 373 GGTGAGGCGGTTCAGGCGGAGGTGGCTCTGGCGGTGGAGATCTCAGGCTGTGGACA 432
Oy 140 GlnSerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLysAla 159
Db 433 CAG--GAATCTGACATCAACCATCACCCTGGTGAAGACACACTACCTTGCTGCTCA 489
Oy 160 SerGlnAspValTyr-----AsnAlaValAlaTyrTyrGlnGlnLysProGlyGln 176
Db 490 AGTACTGGGGCTGTACAACACTAGTACTATGCGCAATGGGCTCCAAAGAAACCAATCAT 549
Oy 177 SerProLysLeuLeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPhe 196
Db 550 TTATTCACTGGCTTAAATAGTGTGACCAACACGAGCTCCAGGTTCTCTCCGAGATT 609
Oy 197 ThrGlySerGlySerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAsp 216
Db 610 TCAGGCTCCGATGATGGAGACAGAGCTCCCTCACATCACAGGAGGAGCTGAGGAT 669
Oy 217 LeuAlaValTyrPheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyThr 236
Db 670 GAGGCAATATATTTCTGTCTTATGATGACAGCAACCATTTGGGTGGTGGAGAGAAC 729
Oy 237 LysLeuGlnIleLysValAlaLeuGlnIleSerAsnSerValMetTyrPheSerSerValVal 256
Db 730 AAATCTGACTCTC-----CTCGAG----- 747
Oy 257 ProValLeuGlnLysValAsnSerThrThrLysProValLeuArgThrProSerPro 276
Db 748 -----GAATTCACTACTACCAAGCCAGTCTCGCAATCCCTCACCT 789
Oy 277 ValHisProThrGlyThrSerGlnProGlnArgProGlnAspCysArgProArgGlySer 296
Db 790 GTGACCCCTACCGGAGCATCTCAGCCCAAGAGACCAAGATGTGGCCCCCGGTGGCTCA 849
Oy 297 ValLysGlyThrGlyLysAspPheLeuGlnLysAspProLysLysCysTyrLeuLeuAspGly 316
Db 850 GTGAAGGGGACCGGATTTGACTTC-----GACCCCAACTCTGTACTTGTAGATGGA 903
Oy 317 IleLeuPheIleTyrGlyValIleIleThrAlaLeuTyrLeuArgAlaLysPheSerArg 336
Db 904 ATCTCTTCATCTACGAGATCATCATCAGCCCTGACTGAGACGAAATTTCAGCAGG 963
Oy 337 SerAlaGluThrAlaAlaAsnLeuGlnAspProAsnGlnLeuTyrAsnGlnLeuAsnLeu 356
Db 964 AGTCAGAGACTGCTGCGCAACCTGAGAGACCCCAACGAGCTCTTAAATAGACTCATATTA 1023
Oy 357 GlyArgArgGlnGluTyrAspValLeuGlnLysLysArgAlaArgAspProGlnLysGly 376

```

```

Db 1024 GGGCGAAGAGGAGATATGAGCTCTTGAGAAAGACCGGCTCGGGATCCAGAGATGGGA 1083
Oy 377 GlyLysGlnGlnArgArgArgAspProGlnGlnGlyValTyrAsnAlaLeuGlnLysAsp 396
Db 1084 GCGAAACAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1143
Oy 397 LysMetAlaGluAlaTyrSerGlyIleGlyThrLysGlyGlnArgArgGlyLysGly 416
Db 1144 AACATGCGAAGAGCTTACAGTGAAGATCGGCACAAAGCCGAGAGCGGAGAGAGAGAG 1203
Oy 417 HisAspGlyLeuTyrGlnGlyLeuSerThrAlaThrLysAspThrTyrAspAlaLeuHis 436
Db 1204 CACGATGCGCTTTACAGAGGTCTCAGCACTGCCACCAAGAGACACCTATGATGCCCTGAT 1263
Oy 437 MetGlnThrLeuAlaProArg 443
Db 1264 ATGACAGACCTTGCGCCCTCGC 1284

RESULT 3
A93936 3058 bp DNA linear PAT 22-JAN-2000
LOCUS Sequence 1 from Patent WO9720938.
ACCESSION A93936
VERSION A93936.1 GI.6742038
KEYWORDS
SOURCE
ORGANISM
unidentified
unclassified.
REFERENCE
1 (bases 1 to 3058)
AUTHORS
Rosenthal, F. and Kulmburg, P.
TITLE
CELLS WITH HYBRID RECEPTOR AND GENE CONSTRUCT WHICH CAN BE
CONTROLLED BY SAID HYBRID RECEPTOR, AND USE OF SAID CELLS IN GENE
THERAPY
PATENT: WO 9720938-A 1 12-JUN-1997;
JOURNAL
ROSENTHAL, FELICIA (DE); KULMBURG, PETER (DE)
FEATURES
source
1..3058
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
ORIGIN
Alignment Scores:
Pred. No.: 1,21e-107 Length: 3058
Score: 1501.50 Matches: 297
Percent Similarity: 76.29% Conservative: 44
Best Local Similarity: 66.44% Mismatches: 79
Query Match: 64.03% Indels: 27
Gaps: 6
US-09-596-774-7 (1-443) x A93936 (1-3058):
Oy 1 GlnIleGlnLeuValGlnSerGlyProGlnLeuLysProGlyValThrValLysIle 20
Db 862 CAGGTCCAACTGACGAGCTCTGGGGCTGAGCTTGGAAGCTCGGGGCTCACTGAAGCTG 921
Oy 21 SerGlyValAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTrpValLysGlnAla 40
Db 922 TCCTGCAAGGCTTCGGCTACACCTTCACACCTGATGCTGGGTGAAGCAGAGG 981
Oy 41 ProGlyGlnGlyLeuLysTrpMetGlyTyrIleAsnThrSerThrGlyLysSerThrPhe 60
Db 982 CTGGACGAGGCGCTTGAAGTGGAGATTGATCCATATAGTGGTGTACTAAGTAC 1041
Oy 61 AlaAspAspPheLysGlyArgPheAspPheSerLeuGlnThrSerAlaAsnThrAlaTyr 80
Db 1042 AATGAGAAGTTCAAGACGAGGCGCACACTGACTAGACAAACCTCCAGACAGCCTTAC 1101
Oy 81 LeuGlnIleAsnAsnLeuLysSerGlnAspMetAlaThrTyrPheCysAlaArgTrpGlu 100
Db 1102 ATGACGCTCAGACGCTGACATCTGAGACTCTCGGCTATATTATTTGCAAGATACGAT 1161

```

```

QY 101 ValTYR---HisGlyTYrValProTYrTPrgLYngLYThrThrValThrsSer 119
DB 1162 TACTACGGAGTAGAGTACTTTGACTGAGGCCCAAGGACACCGGTCACTCTCCCTCA 1221
QY 120 GYGLYGLYGLYSerGlyGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThr 139
DB 1222 GGTGAGGCGGCTTCAGGCGAGGTGGCTCTGGCGGTGGCATCTCAGGCTTTGGGACA 1281
QY 140 GlnSerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLeuAla 159
DB 1282 CAG---GAATCTGCACATACCATACCTGGTGAAGAGTCACCTACTGTCGCTCA 1338
QY 160 SerGlnAspValTYR-----AsnAlaValAlaTPTyrGlnGlnLysProGlyGln 176
DB 1339 AGTACTGGGGCTGTACAACACTAGTAACCTATGCACTGGGTCCAGAAAGAAACCATAT 1398
QY 177 SerProLysLeuLeuIleTYrSerAlaSerSerArgTYrThrGlyValProSerArgPhe 196
DB 1399 TTATTCAGTGGCTGTAATAGTGGTGTACCAACACCAAGCTCCAGGTGTCTCCGAGATT 1458
QY 197 ThrGlySerGlySerGlyProAspPheThrPheThrIleSerSerValGlnAlaGlnAsp 216
DB 1459 TCAGGCTCCCTGATGGAGACAGGCTGCCTCCATCACAAGGCGCACAGACTGGGAT 1518
QY 217 LeuAlaValTYrPheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyThr 236
DB 1519 GAGGCAATATATTTCTGTCTGTATGCTATGTAACGACCAATCATGGGTGCGTGGAGGAAC 1578
QY 237 LysLeuGlnIleLysAlaLeuGlnIleSerAsnSerValMetTYrPheSerSerValVal 256
DB 1579 AAACGACTGTC-----CTCGAG----- 1596
QY 257 ProValLeuGlnLysValAsnSerThrThrLysProValLeuArgThrProSerPro 276
DB 1597 -----GAATTCACCTACTCTCCAAAGCCAGTCTGCGAATCTCCCTCACT 1638
QY 277 ValHisProThrGlyThrSerGlnProGlnArgProGlnAspCysArgProArgGlySer 296
DB 1639 GTGCACCTTACCGGAGCATCTCAGCCCAAGACAGAAAGATTGCGGCCCGGCTCA 1698
QY 297 ValLysGlyThrGlyLysAspPheLeuGlnAspProLysLeuCysTYrLeuLeuAspGly 316
DB 1699 GTGAGGAGGACCGGATGGACTTC-----GACCCCAAACTCTCTGCTGATGCA 1752
QY 317 IleLeuPheIleTYrGlyValIleIleThrAlaLeuTYrLeuArgAlaLysPheSerArg 336
DB 1753 ATCTCTTCACTACGAGTATCATCATCAGCCCTGTACTGAGACGAAATTCAGACAG 1812
QY 337 SerAlaGlnThrAlaAlaAsnLeuGlnAspProAsnGlnLeuTYrAsnGlnLeuAsnLeu 356
DB 1813 AGTCACAGAGACTGCTGCCCAACTGCAAGACCCCAACGCTCTACATAGAGCTCAATCTA 1872
QY 357 GlyArgArgGlnGlnTYrAspValLeuGlnLysLysArgAlaArgAspProGlnLysGly 376
DB 1873 GGGGGAAGAGAGAAATGACGTCTTGAGAAAGAGGCGGTGCGGATCCAGAAATGGA 1932
QY 377 GlnLysGlnGlnArgArgAspProGlnGlnGlnValTYrAsnAlaLeuGlnLysAsp 396
DB 1933 GCGAAACAGCAGAGAGAGAGAACCCCAAGAGACCATATACATGACCTGCAAGAAAGAC 1992
QY 397 LysMetAlaGlnAlaTYrSerGlnIleGlyThrLysGlyGlnLysArgArgGlyLysGly 416
DB 1993 AACATGCGACAGAGCCTACAGTGAATCGGCAAAAAGCGAGAGCGAGAGCGCAAGGG 2052
QY 417 HisAspGlyLeuTYrGlnGlyLeuSerThrAlaThrLysAspThrTYrAspAlaLeuHis 436
DB 2053 CACATGAGCCTTACAGAGGCTCAGACATGCGCACCAAGAGACCTATATGATGCGCTGCAT 2112
QY 437 MetGlnThrLeuAlaProArg 443
DB 2113 ATGCAGACCTCGGCGCTTCGC 2133

```

RESULT 4

```

A57337
LOCUS A57337 2070 bp DNA linear PAT 03-MAR-1998
DEFINITION Sequence 7 from Patent EP0739984.
ACCESSION A57337
VERSION A57337.1 GI:3713216
KEYWORDS
SOURCE
ORGANISM
synthetic construct
other sequences/ artificial sequences.
REFERENCE
1 Wels,W.D., Schmidt,M. and Groner,B.D.
Bivalent polypeptides containing at least two domains
Patent: EP 0739984-A 7 30-OCT-1996;
SAN TUMORFORSCHUNGS GMBH (DE)
Location/Qualifiers
1..2070
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/clone="pSW202-5"
1..1956
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAA03441.1"
/db_xref="GI:3713217"
/translation="MKKTALIAVALAGFATVAQADYKDDDKLHHHHHKKLOVOLOQ
SGPELKKPGETVAKISKASGYFTNTGMMVVKQAPQGLKMGWINTSGESTFADF
KGRFDFSLSTANTAYLQINLKSEDMATYFCARMVHYGVYVWQGTGVSVSSG
GGGGGGGGSDIQLTQSHKELSTGSRVSLTCKASODVNAVMYQKQSGSKPL
LIYSASRYTGVPSRFTGSGSGDPFTTSSVOAEDLAVYPCQHPRTFTGSGTKL
EIKALSHHHHHLEGGSLAALTAHQAQCHLPLEFTTHRQPRGMEOLECGGYVORIVA
LYAARLSNOUNDOVTRNMLASPGSGDGEAIRBPBQARIALTLIAAESRFPQSG
TGNDERGANADVSLITCPVALGECAPGSDSGDLUBERYPTGAELFGSDVSESTR
GTQNWTVERLLQHRQLBERGYFVYGHTFLEAQSIVFGVRAISODLDAWRGFY
IADDPALAYGADQDEPDARGRIRNGALIRVYVPSLPGFYRTSLTLAAPRAGFE
RLIGHPLPLRLDITGPEEGGRLFTILGMPLAERTVIPSALPTDPRNVGDLDPSS
IPKEXGALSLPDPVAGPKPFRBDLK"
ORIGIN
Alignment Scores:
Pred. No.: 4,56e-93 Length: 2070
Score: 1312.50 Matches: 284
Percent Similarity: 68.59% Conservative: 13
Best Local Similarity: 65.59% Mismatches: 55
Query Match: 55.97% Indels: 83
DB: 6 Gaps: 8
US-09-596-774-7 (1-443) x A57337 (1-2070)
QY 1 GlnIleGlnLeuValGlnSerGlyProGlnLysLysPheGlyGlnThrValLysIle 20
DB 115 CAGGTACAACTGACAGCAGTCTGGAACCTGAACAGAACCTGGAGACAGCAAGATC 174
QY 21 SerCysLysAlaSerGlyTYrProPheThrAsnTYrGlyMetAsnTrpValLysGlnAla 40
DB 175 TCTGCAAGGCGCTCTGGATCTCTTCAACAACATAGATGATGAAGTGGTGAAGCAGCT 234
QY 41 ProGlyGlnGlyLeuLysTrpMetGlyTYrPileAsnThrSerThrGlyGlnSerThrPhe 60
DB 235 CCAGGACAGAGGTTTAAAGTGATGGGCTGATTAACCTCCACCTGAGAGATCAACATT 294
QY 61 AlaAspAspPheLysGlyArgPheAspPheSerLeuGlnThrSerAlaAsnTrpAlaTYr 80
DB 295 GCTGATGACTTCAAGGAGCGGTTGACTTCTTTGAAACCTTGCCAAACCTGCTAT 354
QY 81 LeuGlnIleAsnAsnLeuLysSerGlnAspMetAlaThrTYrPheCysAlaArgTrpGln 100
DB 355 TTGCAGATCAACAACCTCAAAAGTGAAGCATGGCTACATATTCTGTGCAAGATGGAG 414
QY 101 ValTYrHisGlyTYrValProTYrTPrgLYngLYThrThrValThrsSerGly 120
DB 415 GTTACCAAGGCTTACTGAGGCGCAAGGAGACCAAGGATCACCGCTTCCTCTGCG 474

```

```

QY 121 GlyGlyGlySerGlyGlyGlySerGlyGlyGlySerAsp116GlnLeuThrGln 140
DB 475 GGTGGCGGTTCTGGTGGGGGTGGCTCCGGCGGTGGCTTGCATCCAGCTGACCCAG 534
QY 141 SerHisGlyPheLeuSerHisSerValGlyAspArgValSer116ThrCysAlaSer 160
DB 535 TCTCAAAATTCCTCTCCACTTCAGTAGAGAGACAGGCTCAGCATCCTGCAAGGCCAGT 594
QY 161 GlnAspVal117AsnAlaVal118ProPyrGln119ProGlyGlnSerProGlyLeu 180
DB 595 CAGAGTGTATTAATGCTGTGTCTCTGTATCAACAGAAACAGACAACTCTTAACCTT 654
QY 181 LeuLeuSerAlaSerSerSerArgTyrThrGlyVal119ProSerArgPheThrGlySerGly 200
DB 655 CTGATTACTCGGCATCTCTCCCGGTACACTGAGTCCCTTCCTGCTTCACTGGCAGTGGC 714
QY 201 SerGlyProAspPheThrPheThr119SerSerVal119GlnAlaGluAspLeuAlaVal119 220
DB 715 TCTGGGCGGATTTCACTTTCACATTCAGACAGTGTGAGAGCTGAGACCTGGCAGTTAT 774
QY 221 PheCysGlnGln118PheArgThrProPheThrPheGlySerGlyThr119LeuGln119 240
DB 775 TTTCTGTCAAGCAATTTCTGTACTCAATTCAGTTCGGCTCGGAGCAAAATTTGAGATC 834
QY 241 LysAlaLeuGlu119SerAsnSerVal119MetTyrPheSerSerVal119ProVal119Gln 260
DB 835 AAAGCTTAGACCA----- 848
QY 261 LysValAsnSerThrThrThrLysProVal119LeuArgThrProSerProVal119ProThr 280
DB 849 -----CCATCATCAACATCACTAGA 869
QY 281 GlyThrSerGlnProGlnArgProGluAspCysArgProArgGlySerVal119GlyThr 300
DB 870 GGG---CGGAGGCTGGCGCGGTGAC-----CGGAGCAACAGGCTGCCACTGCCCT 920
QY 301 GlyLeuAspPheLeuGluAspProLysLeuCysTyrLeuLeuAspGly119LeuPhe119 320
DB 921 GGAAGCTTTCACCCCTCATCGCCACCGCGGCTGGGAGCAACATCGAGCAGTGGCGCTA 980
QY 321 TyrGlyVal119LeuThrAlaLeuTyrLeuArgAla119PheSer----- 335
DB 981 TCCGGTCAAGCGGCTGTGCTCTTCACTGCGCGGCACTGATGGAACAGTGCAG 1040
QY 336 -----ArgSerAlaGluThrAla-----AlaAs 343
DB 1041 CCAAGTGATCCGCAACGCCCTGGCCAGCCCGGCAAGCGGCACTGGGCGAAGCAT 1100
QY 343 nLeuGlnAspProAsnGlnLeuTyrAsnGluLeuAsnLeuGlyArgArgGluGluTyrAs 363
DB 1101 CCGGAGACGCGGAGCGAGCCGCTGTGACCTT-GGCCCGCGCGAGAG----- 1154
QY 363 pValLeuGlnLysArgAlaArgAspProGluMetGlyGlyGlnGlnArgArg 383
DB 1155 -----CGAGGCTTCCGTCCG-----GCAAGGCAACCGGCAACGAGCA-- 1190
QY 383 GAsnProGlnGluGlyVal119AsnAlaLeuGlnLysAspLysMetAlaGluAlaTyrSe 403
DB 1190 ----- 1190
QY 403 rGluLeuGlyThrLysGlyGluArgArgGlyLys 415
DB 1191 -----GGCCGGCGGCGCAACCGCGAGTGTAG 1220

```

```

RESULT 5
LOCUS A57353 2730 bp DNA linear PAT 03-MAR-1998
DEFINITION Sequence 23 from Patent EP073984.
ACCESSION A57353
VERSION A57353.1 GI:3713232
KEYWORDS Synthetic construct

```

```

ORGANISM synthetic construct
REFERENCE other sequences; artificial sequences.
AUTHORS 1
TITLE Wells W.D., Schmidt M. and Groner B.D.
JOURNAL Bivalent polypeptides containing at least two domains
Patent: EP 073984-A 23 30-Oct-1996;
SAN TUDORFORSCHUNGS GMBH (DE)
FEATURES
source
1..2730
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/clone="PMS246-5-5 KDEL"
1..2700
/notes="unnamed protein product"
/codon_start=1
/transl_table=1
/protein_id="CA03449.1"
/db_xref="GI:371323"
/translation="MKKTAIAIALAGFATVAADYKDDDKLHHHHHKLQVQLQ
SGPELKKRGERTVTKSCKASGYPFTNYGMNVKQAPGGGLKMGWINTSGSTPADP
KGRDPSELETSANLTIQINLKSEDMATYRCAMREYHGVPTVGGCTTYVSSGG
GSGGGSGGGSDDIQLTQSHKFLSTVSDRVSTICKASQDVNVAVMTOQKPGSPKL
LIYASASRYTGVPRTGSGSDPFTFTISSVQAEADLAVPCCQHFRTPTGSGTKL
EIKALEHHHHHLEGSLLAALTAHQACHLPLETPTPRHQPGRMEQLCECGPVORLVA
LYLAARLSMNVQDVIRNALASPGSGDLGAIIEOPQARLALTLLAAESERFVROG
TGNDAGAAANDVSLTCCPVACGACGAPDGDLLERNVPTGABFGDGDVPSSTR
GTQNWYVBRLLQARQLREGRYVFGHTGLTEAASQIVFGVYARASQDDLAIRGFY
IAGDPALAYGAQDQEPDARGIRNGALIRYVPRSSLPGRYRSLTTLAEPALAEVE
RLIGHPLRLADATGPEBEGRLETLIGMLARVTVPALPTDPNNVGGDLDPSS
IPDKXQALISALPDVASQPKPRPASQVQLOQSGPELKPPEYTKISCKASGYPTNY
GMNVKQAPGGGLKMGWINTSGSTPADPDKFRTPTSETSANTAYLAQINLKSE
MATYRCAMREYHGVPTVGGCTTYVSSGGSGGSGGGSDDIQLTQSHKFLST
VGDRVSTICKASQDVNVAVMTOQKPGSPKLIYASASRYTGVPRTGSGSDPFT
FTISSVQAEADLAVPCCQHFRTPTFTSGTKLEIKALEKDEL"
ORIGIN
Alignment Scores:
Pred. No.: 6,44e-93 Length: 2730
Score: 1312.50 Matches: 284
Percent Similarity: 68.59% Conservative: 13
Best Local Similarity: 65.59% Mismatches: 55
Query Match: 55.97% Indels: 83
DB: 6 Gaps: 8
US-09-596-774-7 (1-443) x A57353 (1-2730)
QY 1 Gln119GlnLeuVal119GlnSerGlyProGluLeuLysProGlyGluThrVal119 20
DB 115 CAGGTACAACTGACGACGAGCTGTGACCTGAACTGAAGAACTGAGAGACAGTCAAGTTC 174
QY 21 SerCysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrVal119GlnAla 40
DB 175 TCTGCAAGGCTCTGGGTATCTTTCACAAACTATGAAAGAACTGGGTGAACGAGCT 234
QY 41 ProGlyGlnGlyLeuLysTyrMetGlyTyrTrp119LeuThrSerThrGlyGluSerThrPhe 60
DB 235 CCAGAGACGGGTTTAAAGTGAATGAGCTGATTAACCTTCACTGAGAGATCAACATTT 294
QY 61 AlaAspAspPheLysGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80
DB 295 GCTATATGACTTCAAGGAGACGGTTTGACTTCTTGGAAACCTTGCACAACTGCCAT 354
QY 81 LeuGln119LeuAsnLeuLysSerGlyAspMetAlaThrTyrPheCysAlaArgTyrGlu 100
DB 355 TTGCAGATCAACCACTCAAAAGTGAAGACATGCTCATATTTCTGTGCAAGTGGAG 414
QY 101 Val119Thr119GlyTyrVal119ProTyrTrpGlyGlnGlyThrThrVal119SerSerGly 120
DB 415 GTTACCAAGGCTACGTTCTTACTGGGCGCAAGGACCAAGGTCACCGTTCTCTGGC 474
QY 121 Gly119GlySerGlyGlyGlySerGlyGlyGlySerAsp116GlnLeuThrGln 140

```

```

Db 475 GGTGGCGGTTCTGTGGCGGTGGCTCCGGCGGTGGCTTGACATCCAGTCGACCGAG 534
Qy 141 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSer 160
Db 535 TCTCAAAATTCCTCTCCATCTTCAGTAGAGACAGGGTCAGATCCTCGAAGGCCAGT 594
Qy 161 GlnAspValTyrAsnAlaValAlaIleTyrGlnGlnLysProGlyGlnSerProLysLeu 180
Db 595 CAGAGTGTATTAATGCTGTGGCTCGGTATCAACAGAAACAGGCAATCTCTTAACCTT 654
Qy 181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200
Db 655 CTGATTACTCGGCATCTCTCCGCTACCTGAGTCCCTTCTCGCTTCACTGGCAGTGGC 714
Qy 201 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGlnAspLeuAlaValTyr 220
Db 715 TCTGGGCGGATTTTCATCTTCACTTCACATGAGTGGAGGCTGGAAGACTGAGTTTAT 774
Qy 221 PheCysGlnGlnIlePheArgThrProPheThrPheGlySerGlyThrLysLeuGlnIle 240
Db 775 TTCTGTCAACAACATTTTGTACTTCATTCAGTTCCGCTCGGGGACAAATTTGGAGATC 834
Qy 241 LysAlaLeuGlnIleSerAsnSerValMetTyrPheSerSerValValProValLeuGln 260
Db 835 AAAGCTCTAGAGCA----- 848
Qy 261 LysValAsnSerThrThrThrLysProValLeuArgThrProSerProValIleProThr 280
Db 849 -----CCATCATCACTCACTCACTGA 869
Qy 281 GlyThrSerGlnProGlnArgProGlnAspCysArgProArgLysSerValLysGlyThr 300
Db 870 GGG---CGGCGAGCTGGCGCGCTGAC-----CGCGCACAGGCTGGCCACCTGCCGCT 920
Qy 301 GlyLeuAspPheLeuGlnAspProLysLeuLysCysTyrIleLeuAspGlyIleLeuPheIle 320
Db 921 GGAACATTTTCAACCGCTATCGCCGCGCGCTGGAGAACACTGGAGACAGTCCGGCTA 980
Qy 321 TyrGlyValIleIleThrAlaLeuTyrLysArgAlaLysPheSer----- 335
Db 981 TCCGGTCAAGCGGTGGTGGCTCTTACCTGGCGCGGACTGTCAATGGAACAGGTGCA 1040
Qy 336 -----ArgSerAlaGlnThrAla-----AlaAs 343
Db 1041 CCAAGTATCCGCAACGCCCGCGGACCGCGGACGCGGACCTGGCGCAAGCGAT 1100
Qy 343 nLeuGlnAspProAsnGlnLeuTyrAsnGlnLeuAsnLeuGlyArgArgGlnGlyTyrAs 363
Db 1101 CCGGAGACAGCGGAGACAGCGCGCTGGCCCTGACCT--GGCGCGCGCGAGAG----- 1154
Qy 363 pValLeuGlnLysLysArgAlaArgAspProGlnMetGlyGlyLysGlnGlnArgArg 383
Db 1155 -----CGAGGGCTTGTCTCG-----CGAGGGACCGGACAGACAGCA-- 1190
Qy 383 GAsnProGlnGlnGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaGlnAlaTyrSe 403
Db 1190 ----- 1190
Qy 403 rGlnIleGlyThrLysGlyGlnArgArgArgLys 415
Db 1191 -----GGCGCGCGCGCGCAACCGCATGTGTGAG 1220

```

```

RESULT 6
A22539          2012 bp   DNA       linear   PAT 15-NOV-1994
LOCUS          M.musculus/P.aeruginosa Fv(FRP5)-ETa fusion protein.
DEFINITION
ACCESSION      A22539
VERSION        A22539.1 GI:641549
KEYWORDS
SOURCE         synthetic construct
               other sequences: artificial sequences.
REFERENCE      1 (bases 1 to 1012)

```

```

AUTHORS      Weis, W.S., Hynes, N.E., Harwerth, I.M., Groner, B., Hardman, N. and
              Zwick, J.M.
TITLE        Recombinant antibodies specific for a growth factor receptor
JOURNAL      Patent: EP 0502812-A 10 09-SEP-1992;
              CIBA-GEIGY AG
FEATURES
source       Location/Qualifiers
              1..2012
              /organism="synthetic construct"
              /mol_type="unassigned DNA"
              /db_xref="taxon:32630"
              1..1911
              /product="Fv heavy and light chain variable domain and
              exotoxin A fusion protein"
ORIGIN
Alignment Scores:
Pred. No.:      6,3e-93      Length:      2012
Score:          1310.50      Matches:      279
Percent Similarity: 68.51%    Conservative: 19
Best local Similarity: 64.14%  Mismatches:     43
Query Match:    55.88%       Indels:        95
Db:              6          Gaps:          10
US-09-596-774-7 (1-443) x A22539 (1-2012)

```

```

Qy 1 GlnIleGlnLeuValGlnSerGlyProGlyLeuLysLysProGlyGlnThrValLysIle 20
Db 97 CAGGTACACTGACGACGAGTGTGACCTGAACCTGAAGAGCTGGAGACAGTCAAGATC 156
Qy 21 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValLysGlnAla 40
Db 157 TCTCGAAGGCTCTGAGTATCTTTCACAACTATGATGATGAACCTGGGTGAAGCGCT 216
Qy 41 ProGlyGlnGlyLeuLysTyrPheGlyTyrIleAsnThrSerThrGlyLysThrPhe 60
Db 217 CCAGACAGGGTTTAAAGTGAAGGAGGCTGATTAAACCTCCACCTGAGAGTCAACATTT 276
Qy 61 AlaAspAspPheLysGlyArgPheAspPheSerLeuGlnThrSerAlaAsnThrAlaTyr 80
Db 277 GCTGATGACTTCAGAGGACGGTTTGACTCTCTTGGAAACCTGTGCCAACACTGCTTAT 336
Qy 81 LeuGlnIleAsnAsnLeuLysSerGlyAspMetAlaThrTyrPheCysAlaArgTyrGln 100
Db 337 TTCCAGATCAACAACCTCAAAAGTGAAGCATGGCTACATATTTCTGTCCAAAGATGGAG 396
Qy 101 ValTyrHisGlyTyrValProTyrTyrGlyGlnGlyThrValThrValSerSerGly 120
Db 397 GTTTACACAGGCTACGTTCTTACCTGAGGCGCAAGGACACAGCTACCGTTCTCTGCGC 456
Qy 121 GlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlyGlySerAspIleGlnLeuThrGln 140
Db 457 GGTGGCGGTTCTGTGGCGGCTGCGCGGCTGCGGCTGCTTGCATCTCCAGCTGACCCAG 516
Qy 141 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSer 160
Db 517 TCTCAAAATTCCTGTCCTGCACTTCAGTAGAGACAGGTCAGCATCCTCGCAAGCCAGT 576
Qy 161 GlnAspValTyrAsnAlaValAlaIleTyrGlnGlnLysProGlyGlnSerProLysLeu 180
Db 577 CAGAGTGTATTAATGCTGTGGCTCGGTATCAACAGAAACAGGCAATCTCTTAACCTT 636
Qy 181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200
Db 637 CTGATTACTCGGCATCTCTCCGCTACCTGAGTCCCTTCTCGCTTCACTGGCAGTGGC 696
Qy 201 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGlnAspLeuAlaValTyr 220
Db 697 TCTGGCGCGGATTTTCATCTTCACTTCATGAGTGTGAGAGGCTGGAAGACTGGCAGTTTAT 756
Qy 221 PheCysGlnGlnIlePheArgThrProPheThrPheGlySerGlyThrLysLeuGlnIle 240
Db 757 TTCTGTCAACAACATTTTGTACTTCATTCACGTTGGCTCGGGGACAAATTTGGAGATC 816

```

Qy 241 LysAlaLeuGlu-----IleSerAsnSerValMetTyrPheSerSerValPro 257
Db 817 AAAGCTTAAAGGGGCGGAGCGCTGGCGCGGACCGGACAGGCGCTGCCACTGGCG 876
Qy 258 ValLeuGlnLysValAsnSerThrThrLys----- 268
Db 877 CTG-----GAGACTTTCACCCCTCATCCGACCGCGCGCTGGAGACA 921
Qy 269 -----ProValLeuArgThrProSerProValHisProThrLysThr 282
Db 922 CTGGAGCAGTGGCGGTATCTCGGTGACGGCTGGT-CGCCCTCTA---CTGGCGGCGCG 977
Qy 283 Ser-----GlnProGlnArgProGlnArgCysArgProArgGlySerValLysGlyThr 300
Db 978 ACTGTCATGAAACCAAGTCGACCGACGTATCCGCAACCGCCGCGACCGCGGAGCGG 1037
Qy 301 GlyLeuAspPheLeuGlnAspProLysLeuCysTyrLeuLeuAspGlyIleLeuPheIle 320
Db 1038 CGGCGACCTGGGCGGAGCGATCCG----- 1061
Qy 321 TyrGlyValIleIleThrAlaLeuTyrLeuArgAlaLysPheSerArgSerAlaGluThr 340
Db 1062 ----- 1076
Qy 341 AlaAlaAsnLeuGlnAspProAsnGlnLeuTyrAsnGlnLeuAsnLeuGlyArgArgGlu 360
Db 1077 GGGCCGCTGGCGCCCTGACCT-----GGCGCGCGCGCA 1109
Qy 361 GlyTyrAspValLeuGlnLysValArgAlaArgAspProGluMetGlyLysGlnGln 380
Db 1110 GAG-----CGAGCGCTTCCTCCG-----GCAGGCGACCGCGCA 1142
Qy 381 ArgArgArgAsnProGlnGlnGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaGlu 400
Db 1143 CGAGCA----- 1148
Qy 401 AlaTyrSerGlnIleGlyThrLysGlyGlnArgArgGlyLys 415
Db 1149 -----GGCGGCGCGCGCAACCGCGACCTGGTGAG 1178

RESULT 7
128526 128526 2012 bp DNA linear PAT 06-FEB-1997
LOCUS Sequence 13 from patent US 5571894.
DEFINITION 128526
ACCESSION 128526
VERSION 128526.1 GI:1819302
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 2012)
AUTHORS Weis,W.S., Hyner,N.E., Harweth,I.-M., Groner,B., Hardman,N. and Zwickl,M.
TITLE Recombinant antibodies specific for a growth factor receptor
JOURNAL Patent: US 5571894-A 13 05-NOV-1996;
FEATURES location/Qualifiers
source 1..2012
/organism="unknown"
/mol_type="unassigned DNA"

ORIGIN
Alignment Scores:
Pred. No.: 6.3e-93 Length: 2012
Score: 1310.50 Matches: 279
Percent Similarity: 68.51% Conservative: 19
Best Local Similarity: 64.14% Mismatches: 43
Query Match: 55.88% Indels: 95
DB: Gaps: 10
US-09-596-774-7 (1-443) x 128526 (1-2012)

Qy 1 GlnIleGlnLeuValGlnSerGlyProGlnLeuLysValProGlyGlnThrValLysIle 20
|||||
|||||

Db 97 CAGGTAACTGACAGCTCTGACCTGAAGTGAAGAGCGTGGAGACAGTCAAGATC 156
Qy 21 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValLysGlnAla 40
Db 157 TCTGTCAAGAGCGCTCTGGGTATCTTTCACAACTATGAAATGAACTGGGTGAAGAGCT 216
Qy 41 ProGlyGlnGlyLeuLysTyrPheGlyTyrPheAsnThrSerThrGlyLysThrPhe 60
Db 217 CCAGAGACAGGCTTTAAAGTGAATGAGCGCTGATTAACCTCACAATCGAGAGATCAACATT 276
Qy 61 AlaAspAspPheLysGlyArgPheAspPheSerLeuGlnThrSerAlaAsnThrAlaTyr 80
Db 277 GCTATGACTTCAAGAGGAGCGGTGACTTCTTGGAAACCTGTGCAACATCGCCAT 336
Qy 81 LeuGlnIleAsnAsnLeuLysSerGlnAspPheAlaThrTyrPheCysAlaArgTyrGlu 100
Db 337 TTGCAGATCAACAACCTTAAGATGAACATGGCTTACATATTTCTGTGCAGATGGAG 396
Qy 101 ValTyrHisGlyTyrValProTyrTyrPheGlyGlnGlyThrThrValThrValSerSerGly 120
Db 397 GTTTAACACCGGCTACGTTCTTACCTGGGCGCAAGGACACAGTCAACGTTCTCTGGC 456
Qy 121 GlyGlyGlySerGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThrGln 140
Db 457 GGTGCGGCTCTGGTGGCGGTGGCTCCGCGCGGTGGCTTGCATCCAGCTGACCGCAG 516
Qy 141 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSer 160
Db 517 TCTCAACAATTCCTGTCCACTCAGTAGAGACAGGCGACACATCCTGCAAGGCGCT 576
Qy 161 GlnAspValTyrAsnAlaValAlaTyrGlnGlnLysProGlyGlnSerProLysLeu 180
Db 577 CAGATGTGATTAATGCTGTTGCTGCTGATCAACAGAAACAGGACATCTCTAACTT 636
Qy 181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200
Db 637 CTGATTTAATCTGGCATCTCCCGGTACACTGAGATCCCTTCTCTTCACTGCGAGTGGC 696
Qy 201 SerGlyProAspPheThrPheThrIleSerSerValGlnAspLeuAlaValTyr 220
Db 697 TCTGGCGCGATTTCACTTCCATCCATGACAGCTGACGCTGAAAGACTGGCACTTAT 756
Qy 221 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyThrLysLeuGlnIle 240
Db 757 TTCTGTCAAGCAACATTTTCGATCTCATTCACGTTCCGCTGGGAGCAAAATTGGAGATC 816
Qy 241 LysAlaLeuGlu-----IleSerAsnSerValMetTyrPheSerSerValPro 257
Db 817 AAAGCTTAAAGGGGCGGAGCGCTGGCGCGGACCGGACAGGCGCTGCCACTGGCG 876
Qy 258 ValLeuGlnLysValAsnSerThrThrLys----- 268
Db 877 CTG-----GAGACTTTCACCCGTCATCCGACCGCGCGCTGGAGACA 921
Qy 269 -----ProValLeuArgThrProSerProValHisProThrLysThr 282
Db 922 CTGGAGCAGTGGCGGTATCTCGGTGACGGCTGGT-CGCCCTCTA---CTGGCGGCGCG 977
Qy 283 Ser-----GlnProGlnArgProGlnArgCysArgProArgGlySerValLysGlyThr 300
Db 978 ACTGTCATGAAACCAAGTCGACCGACGTATCCGCAACCGCCGCGACCGCGGAGCGG 1037
Qy 301 GlyLeuAspPheLeuGlnAspProLysLeuCysTyrLeuLeuAspGlyIleLeuPheIle 320
Db 1038 CGGCGACCTGGGCGGAGCGATCCG----- 1061
Qy 321 TyrGlyValIleIleThrAlaLeuTyrLeuArgAlaLysPheSerArgSerAlaGluThr 340
Db 1062 ----- 1076
Qy 341 AlaAlaAsnLeuGlnAspProAsnGlnLeuTyrAsnGlnLeuAsnLeuGlyArgArgGlu 360
Db 1077 GGGCCGCTGGCGCCCTGACCT-----GGCGCGCGCGCA 1109

QY 361 GIUTYRAAPVALLEUQILULYELYSARGAALARGAASPProGluMeGlyGlyLeuGln 380
DB 1110 GAG-----CGAGGGCTTCGTCG-----GCAGGGACCGGCA 1142
QY 381 ARGATGAGANPProGlnGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaGlu 400
DB 1143 CGACGA----- 1148
QY 401 ALATYRserGluIleGlyThrLysGlyGluArgArgGlyLys 415
DB 1149 -----GCCCGCGCGGCGACCGCACGCGTGGAG 1178
RESULT 8
A50995 1862 bp DNA linear PAT 07-MAR-1997
LOCUS DEFINITION Sequence 36 from Patent WO9613599.
ACCESSION A50995
VERSION A50995.1 GI:2303794
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1 (bases 1 to 1862)
AUTHORS Wels,W. and Fominaya,J.
TITLE NUCLEIC ACID TRANSFER SYSTEM
JOURNAL Patent: WO 9613599-A 36 09-MAY-1996;
WELLS WINFIELD (DE)
COMMENT Other publication AU 3926895 960523.
FEATURES
source
1. .1862
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
1. .1854
/note="unnamed protein product"
/codon_start=1
/protein_id="CAA03250.1"
/translation="MDYKDDDDKKLHHHHHKLSTLEQACDLCRLKKLCKSKKPK
AKCLKNWECGYSPTKRSPLTRALHLEVSRLERLEOLPLLPREDLMILKMSL
ODIKALLTGLPYQDNVANKDAVTRDLASVEDTMDPTLROHLSATSSSESSNKGRDL
TVSSSLVAGSLSCINDMDVTRDKTKTLESLKEHPINKKSESSEKATVSEKAO
YLEEFHQTALEHPELSLKTVTGTPVPAAGAVAAVNAOVAVIDSTANLEKTVAA
LSILPFGVWGIVADGAVHNTEIEIVASIALSVAAQAPLVGELVIDIGFAVNF
ESIILFOVHNSYNRPAYSGVDGIDKLOVLOQSGPELKKPEETVKSICXSGYF
TNYGMNVKQAPQGLKMGKMGINTSGESTPADDFKGRDPSLETSAATVLIQNNLK
SEDMATYFCARMEYHGYVPTMGQTTVTYSSGGGSGGSGGSGGSDIQLTQSHKFL
STISGDRVSTICKASQDVNANVMTQKPGQSPKLLIYSASRTGVPSPFTGSGGP
DFTFTTSSVQAEIDLAVYFCQHPFTPTFGSGTGLTETKALDELSSERRFSA"

ORIGIN
Alignment Scores:
Pred. No.: 8,98e-93 Length: 1862
Score: 1308.00 Matches: 245
Percent Similarity: 97.64% Conservative: 3
Best Local Similarity: 96.46% Mismatches: 6
Query Match: 55.78% Indels: 0
DB: 6 Gaps: 0
US-09-596-774-7 (1-443) x A50995 (1-1862)

QY 1 GlnIleGlnLeuValGlnSerGlyProGluLeuLysLysProGlyGluThrValLysIle 20
DB 1090 CAGGTACACCTGCGACGACTCTGACCTGAAGTGAAGCCTGGAGAGACAGTCAGATC 1149
QY 21 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValLysGlnAla 40
DB 1150 TCTGCAAGGCTCTGGGTAATCTTTCACAAACTATGAACTGAACCTGGTGAAGAGGCT 1209
QY 41 ProGlyGlnGlyLeuLysTyrPheGlyTyrTrpIleAsnThrSerThrGlyGluSerThrPhe 60
DB 1210 CCAGGACAGGGCTTTAAAGTGAATGGCTGAGATTACACCTCCACTGGAGAGTCAACATTT 1269

QY 61 AlaAspAspPheLysGlyArgPheAspPheSerIleuGlnThrSerAlaAsnThrAlaTyr 80
DB 1270 GCTATGATCTTCANAGGACCGATTGACTCTCTCTTGGAAACCTCTGCCAACACTGCTTAT 1329
QY 81 LeuGlnIleAsnAsnLeuLysSerGluAspMetAlaThrTyrPheCysAlaArgTyrGlu 100
DB 1330 TTGCAGATCAACAACCTCCAAAGTGAAGCATGGCTCATATTCTGTGCAGATGGGAG 1389
QY 101 ValTyrHisGlyTyrValProTyrTyrTrpGlyGlnGlyThrThrValLysSerGly 120
DB 1390 GTTATCCACGCGTACGTTCTTACTGTGCGCCAGGAGCCACCGGTACCGCTTCTCTGCG 1449
QY 121 GlyGlyGlySerGlyGlyGlySerGlyGlyGlyGlySerAspIleGlnLeuThrGln 140
DB 1450 GGTGGCGGCTCTGTGGCGGTGGCTCCGCGGTGGCGGTCTTACATCCAGTCAACCGAG 1509
QY 141 SerHisLysPheLeuSerThrThrSerValGlyAspArgValSerIleThrCysLysAlaSer 160
DB 1510 TCTCACAAATTCCTGTCTCAGTACAGTACAGGAGTCAAGGATCACTGCAAGGCGAGT 1569
QY 161 GlnAspValTyrAsnAlaValAlaTyrTyrGlnGlnLysProGlyGlnSerProLysLeu 180
DB 1570 CAGATGTGTATATGCTGTGTGCTGTGTCAACAGAAACAGAGCAATCTCTTAAACTT 1629
QY 181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200
DB 1630 CTGATTTACTCGGACATCTCCCGGTACACTGAGTCCCTTCTGCTTCACTGCAAGTGGC 1689
QY 201 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 220
DB 1690 TCTGGCGCGGATTTCACTTTCACCATCAACAGTGTGCGAGCTGAAGACTGGCAGTTTAT 1749
QY 221 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyThrLysLeuGluIle 240
DB 1750 TTCTGTACGAAATTTTGCTACTTCATTCACCTTGGCTGGGCAAAATTTGAGATC 1809
QY 241 LysAlaLeuGluIleSerAsnSerValMetTyrPheSerSer 254
DB 1810 AAAGCTCTAGAGATCTCTCGATGAGACAAAGTTTTCAGCC 1851
RESULT 9
AR268367 1862 bp DNA linear PAT 10-APR-2003
LOCUS AR268367
DEFINITION Sequence 36 from patent US 6498233.
ACCESSION AR268367
VERSION AR268367.1 GI:29698717
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1862)
AUTHORS Wels,W. and Fominaya,J.
TITLE Nucleic acid transfer system
JOURNAL Patent: US 6498233-A 36 24-DEC-2002;
FEATURES
source
1. .1862
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Alignment Scores:
Pred. No.: 8,98e-93 Length: 1862
Score: 1308.00 Matches: 245
Percent Similarity: 97.64% Conservative: 3
Best Local Similarity: 96.46% Mismatches: 6
Query Match: 55.78% Indels: 0
DB: 6 Gaps: 0
US-09-596-774-7 (1-443) x AR268367 (1-1862)

QY 1 GlnIleGlnLeuValGlnSerGlyProGluLeuLysLysProGlyGluThrValLysIle 20

```

Db      1090 CAGGTACAACTGACAGCTGAGCTGGAACCTGAAAGAAAGCTGAGAGACACTCAAGATC 1149
Qy      21  SerCysValAspAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTrpValIleGlnIla 40
Db      1150 TCTGCAAAAGGCTCTGGGTATCTCTTCAAACTAGATGATACCTGGGTGAAGCAGGCT 1209
Qy      41  ProGlyGlnGlyLeuIleYsrTrpMetGlyTrpIleAsnThrSerThcGlyIleSerThrPhe 60
Db      1210 CCAGGACAGGGTTTAAAGTCGATGGGGCTGATTAACTCCACCTCAGCGAGAGCAACATTT 1269
Qy      61  AlaAspAspPheIleYsrGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80
Db      1270 GCTGATGACTTCAAGGGAAGCGGTTTACCTCTTTGGAAACCTCGCAACACTCTCCAT 1329
Qy      81  LeuGlnIleAsnAsnLeuIleYsrSerGluAspMetAlaThrTyrPheCysAlaArgTrpGlu 100
Db      1330 TTGCAGATCAACCACTCAAAAGTAGAAGACATGCTACATATTTGTGCAAGATGGAG 1389
Qy      101 ValTyrHisGlyTyrValProTyrTrpGlyGlnGlyThrValThrValSerSerGly 120
Db      1390 GTTTRACACGGCTACGTTCTCTTACCTGGGCCAAGGACCAAGCTACCGTTTCTCTGGC 1449
Qy      121 GlyGlyGlySerGlyGlyGlySerGlyGlyGlySerAspIleGlnIleThrGln 140
Db      1450 GGTGGCGGTTCTGTGGCGGTGGCTCCGGCGGTGGCGGTTCTGACATCCAGCTGACCAG 1509
Qy      141 SerHisIleYsrPheLeuSerThrSerValGlyAspArgValSerIleThrCysValAlaSer 160
Db      1510 TCTCAAAATTCCTGTCACTTCACTAGAGAGACAGGTCAGCATCACTGCAAGGCCAGT 1569
Qy      161 GlnAspValTyrAsnAlaValAlaTrpTyrGlnGlnIleYsrProGlyIleSerProIleLeu 180
Db      1570 CAGATGCTGTATATGCTGTTGCTGCTGATCAACGAAACCAAGCAATCTCTTAATT 1629
Qy      181 LeuIleYsrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200
Db      1630 CTGATTTACTCGGCATCTCCCGGATACACTGAGATGCCCTTCTCGCTTCACTGGCAGTGGC 1689
Qy      201 SerGlyProAspPheThrPheThrIleSerSerValGlnIleGluAspLeuAlaValTyr 220
Db      1690 TCTGGGCGGATTTCATCTTCACTCACTGACAGCATGTCAGGCTGAAAGCCTGCACTTAT 1749
Qy      221 PheCysGlnGlnIleAspPheArgThrProPheThrPheGlySerGlyThrIleLeuGluIle 240
Db      1750 TTCTGTACGACAACTTTTCGTTCTCATTCATTCAGCTTCGGCTGGGGAACAAATTTGAGATC 1809
Qy      241 LysAlaLeuGlnIleSerAsnSerValMetTyrPheSerSer 254
Db      1810 AAAGCTTAGAGGATCTCTCGAGTGAAGAAAGATTTTCAGCC 1851

RESULT 10
A50993  A50993 1919 bp DNA linear PAT 07-MAR-1997
LOCUS   Sequence 34 from Patent WO9613599.
DEFINITION A50993
ACCESSION A50993.1 GI:2303792
VERSION   A50993.1 GI:2303792
KEYWORDS
SOURCE    unidentified
           unidentified
           unclassified.
REFERENCE 1 (bases 1 to 1919)
AUTHORS   Weis,W. and Fominaya,J.
TITLE     NUCLEIC ACID TRANSFER SYSTEM
JOURNAL   Patent: WO 9613599-A 34 09-MAY-1996;
          WELLS WINFRIED (DE)
          Other publication AU 3926895 960523.
COMMENT   Location/Qualifiers
FEATURES
          source
            1..1919
              /organism="unidentified"
              /mol_type="unassigned DNA"
              /db_xref="taxon:32644"
              64..1908
                /note="unnamed protein product; Protein sequence is in

```

```

conflict with the conceptual translation"
/codon_start=1
/protein_id="CA003249.1"
/db_xref="gi:2303793"
/translation="DYKDDDDKLLHHHHHKLSSIEQACDLCRLKLCSEKREKCAK
CLKNWEKCRVSPKTRKSLPLTRALHTEVSRERLEQLPLIFPRDLMLIKMSLOD
IKALITGLFVODNNKDAVDRDLASVEDEMLTLRORHISATSSSESSNGOROLTV
SSSLAVGSSLSICINLMDVIDIKTKTESLKEHPIKMKSESINKTVESEKAKQYL
EEFHOTALEHELSLKVGTGNTVPFAGANYAAVAVAVOYIDSETADNLEKTPALS
ILPISGVMGTLADAVHNTREIYAQSLAISLVAVALPLVGLVDIIGFANVRES
IINLFOVHNSINRPAISPGVDIDKQVQLQSGPEPKKGETYKISCKASGTFPTN
YGNMVAQAPDQGLKMGWINTSTSESTPADPFGKRFDSLETANTAVLYLINDLKS
DMATYFCARMEVGVHYVYVWQCTTVIVSSGGSGGGSGGSDIQLTSSHRLST
SVGRVSTICKASODVYNAVAMYQKFGQSPKLIYASRSRYTGVPSRFTSGSGPDR
TFTISSVAEDLAVYFCQGHRTFTFTGSSGTKEIKLELSSRRRSA"

ORIGIN

Alignment Scores:
Pred. No.: 9,32e-93 Length: 1919
Score: 1308.00 Matches: 245
Percent Similarity: 97.64% Conservative: 3
Best Local Similarity: 96.46% Mismatches: 6
Query Match: 55.78% Indels: 0
DB: 6 Gaps: 0

US-09-596-774-7 (1-443) x A50993 (1-1919)

Qy      1  GlnIleGlnLeuValGlnSerGlyProGlyLeuIleYsrProGlyIleThrValIle 20
Db      1147 CAGGTACAACTGACAGCTGAGCTGGAACCTGAAAGAAAGCTGAGAGACACTCAAGATC 1206
Qy      21  SerCysValAspAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTrpValIleGlnIla 40
Db      1207 TCTGCAAAAGGCTCTGGGTATCTCTTCAAACTAGATGATACCTGGGTGAAGCAGGCT 1266
Qy      41  ProGlyGlnGlyLeuIleYsrTrpMetGlyTrpIleAsnThrSerThcGlyIleSerThrPhe 60
Db      1267 CCAGGACAGGGTTTAAAGTCGATGGGGCTGATTAACTCCACCTCAGCGAGAGCAACATTT 1326
Qy      61  AlaAspAspPheIleYsrGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80
Db      1327 GCTGATGACTTCAAGGGAAGCGGTTTACCTCTTTGGAAACCTCGCAACACTGCTTAT 1386
Qy      81  LeuGlnIleAsnAsnLeuIleYsrSerGluAspMetAlaThrTyrPheCysAlaArgTrpGlu 100
Db      1387 TTGCAGATCAACCACTCAAAAGTAGAAGACATGCTACATATTTCTGCAAGATGGAG 1446
Qy      101 ValTyrHisGlyTyrValProTyrTrpGlyGlnGlyThrValThrValSerSerGly 120
Db      1447 GTTTRACACGGCTACGTTCTCTTACCTGGGCCAAGGACCAAGCTACCGTTTCTCTGGC 1506
Qy      121 GlyGlyGlySerGlyGlyGlySerGlyGlyGlySerAspIleGlnIleThrGln 140
Db      1507 GGTGGCGGTTCTGTGGCGGTGGCTCCGGCGGTGGCGGTTCTGACATCCAGCTGACCAG 1566
Qy      141 SerHisIleYsrPheLeuSerThrSerValGlyAspArgValSerIleThrCysValAlaSer 160
Db      1567 TCTCAAAATTCCTGTCACTTCACTAGAGAGACAGGTCAGCATCACTGCAAGGCCAGT 1626
Qy      161 GlnAspValTyrAsnAlaValAlaTrpTyrGlnGlnIleYsrProGlyIleSerProIleLeu 180
Db      1627 CAGATGCTGTATATGCTGTTGCTGCTGATCAACGAAACCAAGCAATCTCTTAATT 1686
Qy      181 LeuIleYsrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200
Db      1687 CTGATTTACTCGGCATCTCCCGGATACACTGAGATCTCTTCTCCCTTCACTGGGAGTGGC 1746
Qy      201 SerGlyProAspPheThrPheThrIleSerSerValGlnIleGluAspLeuAlaValTyr 220
Db      1747 TCTGGGCGGATTTCATCTTCACTCACTGACAGCTGCGGCTGAAAGCCTGGCAGTTTAT 1806
Qy      221 PheCysGlnGlnIleAspPheArgThrProPheThrPheGlySerGlyThrIleLeuGluIle 240

```


Db 1807 TTCTGTACGACAACTTTTCTGATTCATTCAGTTCGGCTCGGGACAAATTGGAGATC 1866
Qy 241 LysAlaLeuGluIleSerAsnSerValMetTyrPheSerSer 254
Db 1867 AAGCTTAGAGAGATCTCTCGAGTGAAGAAAGATTTCAGACC 1908
RESULT 11
AR268366
LOCUS AR268366
DEFINITION Sequence 34 from patent US 6498233.
ACCESSION AR268366
VERSION AR268366.1 GI:29698716
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1919)
AUTHORS Weis,W. and Fomlnaya,J.
TITLE Nucleic acid transfer system
JOURNAL Patent: US 6498233-A 34 24-DEC-2002;
FEATURES
source 1. 1919
Location/Qualifiers
/organism="unknown"
/mol_type="genomic DNA"
ORIGIN
Alignment Scores:
Pred. No.: 9,32e-93 Length: 1919
Score: 1308.00 Matches: 245
Percent Similarity: 97.64% Conservative: 3
Best Local Similarity: 96.46% Mismatches: 6
Query Match: 55.78% Indels: 0
DB: Gaps: 0
US-09-596-774-7 (1-443) x AR268366 (1-1919)
Qy 1 GlnIleGlnLeuValGlnSerGlyProGluLeuLysProGlyGluThrValLysIle 20
Db 1147 CAGGTACAACTGACGACAGTCTGACCTGAACCTGAAGAAAGCTGAGAGACGTCAAGATC 1266
Qy 21 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTrrValLysGlnAla 40
Db 1207 TCCGACGAGGCTTGGGTATCTTTCAACAACATGGAATGAACTGGGTGAACAGGCT 1266
Qy 41 ProGlyGlnGlyLeuLysTrrPheMetGlyTrrIleAsnThrSerThrGlyLysSerThrPhe 60
Db 1267 CCAGGACAGGGTTTAAAGTGGATGGCTGATTAACCTTCACCTGAGAGGTCAACATTT 1326
Qy 61 AlaAspAspPheLysGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80
Db 1327 GCTGATGACCTTCAAGGACGGTGTGACTTCTTTGGAAACCTCTGCCAACACTGCCTAT 1386
Qy 81 LeuGlnIleAsnLeuLysSerGlyLysAspMetAlaThrTyrPheCysAlaArgTrrGlu 100
Db 1387 TTGAGATCAACAACCTCAAAAGTGAAGACATGGCTCAATTTCTGTGCAAGATGGAG 1446
Qy 101 ValTyrHisGlyTyrValProTyrTrrGlyGlnGlyThrThrValThrValSerSerGly 120
Db 1447 GTTACACACGGCTACGTTCTTACTGGGGCCAAAGGACACAGGCTTCTTCTCTGGC 1506
Qy 121 GlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThrGln 140
Db 1507 GGTGGCGGTTCTGTGGCGGTGGCTCCGGCGGTGGCGGTCTGACATCCAGCTGACCCAG 1566
Qy 141 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysGlyAlaSer 160
Db 1567 TCTTCAAAATTCCTGCTCACCTCAGTGAAGACAGGGTCAACCTCAGACAGGCCAGT 1626
Qy 161 GlnAspValTyrAsnAlaValAlaTrrTrrGlnGlnLysProGlyLysSerProLysLeu 180
Db 1627 CAGAGTGTGTAATAGCTGCTTGGCTGATATCAACAGAAACAGGACATCTCTTAACTT 1886
Qy 181 LeuIleTyrSerAlaSerSerArgTrrThrGlyValProSerArgPheThrGlySerGly 200

Db 1687 CTGATTTACTCGGATCTCCCGGTACACTGAGAGTCCCTCTCGCTTCACTGGCAGTGGC 1746
Qy 201 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 220
Db 1747 TCTGGGCGGATTTCACTTTCACCATCAACGACGTGTGACGCTGAAGACCTGGCAGTTTAT 1806
Qy 221 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyThrLysLeuGluIle 240
Db 1807 TTCTGTACGACAACTTTTCTGACTCCATTCACGTTGCGCTCGGGACAAAAATTGAGATC 1866
Qy 241 LysAlaLeuGluIleSerAsnSerValMetTyrPheSerSer 254
Db 1867 AAGCTTAGAGAGATCTCTCGAGTGAAGAAAGATTTCAGACC 1908
RESULT 12
AR57341
LOCUS AR57341
DEFINITION Sequence 11 from Patent EP0739584.
ACCESSION AR57341
VERSION AR57341.1 GI:3713220
KEYWORDS
SOURCE Synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Weis,W.D., Schmidt,M. and Groner,B.D.
TITLE Bivalent polypeptides containing at least two domains
JOURNAL Patent: EP 0739584-A 11 30-OCT-1996;
FEATURES
source 1. 2214
Location/Qualifiers
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/clone="PMS238-5-TGF"
1. 2100
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAA03443.1"
/db_xref="GI:3713221"
/translation="MKRTAIAIAVALAGPATVAQADYKDDDKLHHHHHKLQVLOQ
SGPELKKPEETVKISCKASGYPFTNNGMWWVOAPEGGLKMGWINTSGESTPADP
KRPDPSLETSANTATYIOMNKSEMAIYVPCAPREYVGVYPYQCGQTTVYVSSGG
GSGGGSGSGGSDIQLTOSKRLSTSVGRVITCAASDPYNAVAAVQKQSGSKL
LTVASASRYTGVPSTGSGSDPFTTISVQADLAVYFCQQRTRPTFGSGTKL
EIKALEHHHHHHLEGGSLAALTAHQACHLPLTRHRORPROMEQLEQGYVORLVA
LYLAARLSNVOYDQVRNALASPGSGDGEAIREQPEQARLALTLAAESERFVVG
TGNDEAGANADEKLGTVVSHNDPDSHTQFCFGTCRFLVQEDKPAVCCHSGYVG
ARCEHADILASLESSAGPADSGDALLERVYPTGAFELADGGVSVSTRGTONWYERL
LQAHROLERGVYFVGHGTPLFAAOSIYFGVRRARSQDLDMRGFTYAGDPALAYG
YAQDDPDAKGRIRNGALIRVYVPSLSLGFRTSLTLAAPAGSVERLIGHPLPLR
LDAITGPEERGLRLETLIGMPLAERTVIPSALPTDPRNVGDLDPSIPDKEQAISA
LDVVASQPKPPREDLK"

ORIGIN
Alignment Scores:
Pred. No.: 1,11e-92 Length: 2214
Score: 1308.00 Matches: 275
Percent Similarity: 71.19% Conservative: 24
Best Local Similarity: 65.48% Mismatches: 59
Query Match: 55.78% Indels: 64
DB: Gaps: 5
US-09-596-774-7 (1-443) x AR57341 (1-2214)
Qy 1 GlnIleGlnLeuValGlnSerGlyProGluLeuLysProGlyGluThrValLysIle 20
Db 115 CAGGTACAACTGACGACAGTCTGACCTGAACCTGAAGAAAGCTGAGAGACGTCAAGATC 174
Qy 21 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTrrValLysGlnAla 40


```

Db      175 TCCTGCAAGCCTCTGGTATCTTTCACAACTAGATGATGATGGGTGAGCAGCT 234
Qy      41 ProGlyGlnGlyLeuLysTrpMetGlyTrpIleAsnThrSerThyGlyGluSerThrPhe 60
Db      235 CCAGGACAGGGGTTTAAAGTGATGGCTGGATTAAACCTTCACATGAGAGCAACATT 294
Qy      61 AlaAspAspPheLysGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80
Db      295 GCTATGATCTTCAGAGGACGGTTTACCTTCTTGGAAACTCTGCCAACACTCCCTAT 354
Qy      81 LeuGlnIleAsnAsnLeuLysSerGluAspMetAlaThrTyrPheCysAlaArgTrpGlu 100
Db      355 TTGCAGATCAACAACCTCAAAAGTAGACATGCTACATATTTCTGAGCAAGAGGAG 414
Qy      101 ValTyrHisGlyTyrValProTyrTrpGlyGlnGlyTrpThrValThiValSerGly 120
Db      415 GTTAAACACGGCTACGTTCTTACTGAGGACCAAGGACACGCTACCGTTCTCTGAC 474
Qy      121 GlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThrGln 140
Db      475 GGTGGCGGTTCTGGTGGCGGTGGCTCCGGCGGTGGCGGTCTGACATCCAGCTGACCAG 534
Qy      141 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSer 160
Db      535 TCTCAACAATTCCTGTCCACTTCACTAGAGAGACAGGCTCAGCATCAGCTGCAAGGCCAGT 594
Qy      161 GlnAspValTyrAsnAlaValAlaTrpTyrGlnGlnLysProGlyGlnSerProLysLeu 180
Db      595 CAGATGTGTATTAAGCTGTGCTGCTGATCAACAGAAACCAAGCAATCTCTTAACCTT 654
Qy      181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200
Db      655 CTGATTTACTGGGATCTCTCCCGGTACACTGAGAGCTCTTCTCGCTTCACTGGCAGTGGC 714
Qy      201 SerGlyProAspPheThrPheThrIleSerSerValGlnIleGluAspLeuAlaValTyr 220
Db      715 TCTGGGCGGATTTCACTTTCACATCAGAGCTGAGCGCTGAGAGACTGGCAGATTAT 774
Qy      221 PheCysGlnGlnIleAspPheArgThrProPheThrPheGlySerGlyThrLysLeuGluIle 240
Db      775 TTCGTGTCGCAACATTTTCATCTCATTCACTGCGCTGGGAGCAAAATTTGAGATTC 834
Qy      241 LysAlaIleLeuGluIleSerAsnSerValMetTyrPheSerSerValIleProValLeuGln 260
Db      835 AAAGCTCTTAGACA----- 848
Qy      261 LysValIleAsnSerThrThrLysProValLeuArgThrProSerProValIleProThr 280
Db      849 -----CATCATCAACATCACTAGA 869
Qy      281 GlyThrSerGlnProGlnArgProGluAspCysArgProArgGlySerValLysGlyTyr 300
Db      870 GGG---CGGCAAGCCTGGCGGCGCTGAC-----CGGCAACAGGCTGCACTGCCCT 920
Qy      301 GlyLeuAspPheLeuGluAspProLysLeuCysTyrLysLeuAspGlyIleLeuPheIle 320
Db      921 GGAAGCTTTCACCCCTCATCCGACCGCGCGCTGGAGAACACTGAGACAGTCCGCGCTA 980
Qy      321 TyrGlyValIle-IleThrAlaLeuTyrLeuArgAlaLysPheSerArgSerAlaGluThr 340
Db      981 TCCGGTGCAGGGCGCTGCTGCGCTTCACTGGCGCGGCACTGCTCA----- 1026
Qy      340 PAlaIleAsnLeuGlnAspProAsnGlnLeuTyrAsnGlnLeuAsnLeuGlyArgArgGly 360
Db      1027 -----TGGAAACAGGTCGACCA-GGT----- 1046
Qy      360 uGluTyrAspValLeuGluLysLysArgAlaArgAlaArgLysMetGlyGlyLysGlnGly 380
Db      1047 -----CATCCGCAAGCCCTCGGCGCAAGCCCGC 1072
Qy      380 nArgArgArgAsnPro-GlnGluGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaG 400
Db      1073 GCAGCGGGGGGACTCTGGGCGAAGGATCCGGAGACAGCCGAGAGCGCCGTCTGGCC 1132

```

```

Qy      400 LuAlaTyrSerGluIleGlyThrLysGlyGluArgArgGlyLysGlyHisAsp 418
Db      1133 TGACCTCGGCGCGCGCGAGAGCGAGCGCTTCTGTCGCGACGGCAGCGCAAGC 1188

RESULT 13
A57343
LOCUS       A57343                2793 bp    DNA        linear    PAT 03-MAR-1998
DEFINITION Sequence 13 from Patent EP0739984.
ACCESSION  A57343
VERSION     A57343.1 GI:3713222
KEYWORDS
SOURCE      synthetic construct
            synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
AUTHORS     Wels,W.D., Schmidt,M. and Groner,B.D.
TITLE       Bivalent polypeptides containing at least two domains
JOURNAL     Patent: EP 0739984-A 13 30-OCT-1996;
            SAN TUMORFORSCHUNGS GMBH (DE)
FEATURES
            location/Qualifiers
            source
            1..2793
               /organism="synthetic construct"
               /mol_type="unassigned DNA"
               /db_xref="taxon:32630"
               /clone="PMS238-5-225"
               1..2679
                  /note="unnamed protein product"
                  /codon_start=1
                  /transl_table=1
                  /protein_id="CAA03444.1"
                  /db_xref="gi:3713223"
                  /translation="MKKTAIAlAVAlAGFAVAAQADYKDDDKLHHHHHKLQVLOO
SGBELKRPGEVVKISCAKSGPTPTNGMNTKQAPGGGLKMGKINHTSGSTPADPR
KGRDFSLKTSNANTAYLIQINLKSEDMATTFCAWEVYHGVPTVWGQTVTVSGG
GSGGGGGGGGGDQLOQSHKFLSTVSDRVSITKSAQDVYNAVAYQKPGGPKL
LIYASASRYGVSPTFGSGSPDFTTISVQMEDLAVFCCQHFPTPFYGGGTGL
EIKALEHHHHHLEGGSLAALTAHOAGHLPLETTRHROPFGMLBOCGPYVORLVA
LYLAARLSMNOVDVIRNALAPSGGDGLIAIEOPBOARLALTAAASERFVROG
TGNDACGANDEKLQVLOESGPGLYVOPSLSITCTVSGFSLTNGVHMYRSPK
GLEMLGVYWGSGNTDYNTPFTSRLSINKDKSKOVFFKMSLQSDNTAIYCANALTY
YDFEYFAWGGTIVTVSSGGGSGGGGGGSDILOQSPVILSVSGERVSFSCRA
SOSIGTINHWYQKRTNGSPRLIKYASESISGIPRSFGSGSDFTLINSVSEDI
ADYVCQNMNMPTPTFGAGTKLEIKALDPRPSAGPADSGDALERNYPGAFELD
GGVYSRSTRQNTVTRILQAHQLEERGVPTVCYCTPLAASQIVFGVTRASD
LDATWRGTYLAGDPLAYGTAQDDEPARGRIRGALLRYTPVPSLSPGFRITLTLA
APEAAGVERLIGPLRLDALTGPBEEGRLTTILGWPLAERTVIVPSAIPTPDRN
VGDLDPSSIPDKQAISALPDVYASQPKPRBDLK"

ORIGIN
Alignment Scores:
Pred. No.:      1,49e-92      Length:      2793
Score:          1308.00      Matches:      275
Percent Similarity: 71.19%      Conservative: 24
Best Local Similarity: 65.48%      Mismatches: 59
Query Match:    55.78%      Indels:      64
Db:              6          Gaps:      5

US-09-596-774-7 (1-443) x A57343 (1-2793)
Qy      1 GlnIleGlnLeuValGlnSerGlyProGluLeuLysLysProGlyGluThrValLysIle 20
Db      115 CAGGTACACTGCAGCACTCTGGACCTGACCTGAAGAAGCCTTGAGAGACAGTCAAGTTC 174
Qy      21 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTrpValLysGlnAla 40
Db      175 TCCTGCAAGGCTCTGGGTATCTTTCACAAACTATGGAATGAACTGGGTGAAACAGGCT 234
Qy      41 ProGlyGlnGlyLeuLysTrpMetGlyTrpIleAsnThrSerThyGlyGluSerThrPhe 60
Db      235 CCAGGACAGGGGTTTAAAGTGATGGCTGGATTAAACCTTCACATGAGAGCAACATT 294
Qy      61 AlaAspAspPheLysGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80

```

```

Db      295 GCTGATGACTTCACAGGAGCGGTTGACTTCTTGGAAACCTCGCCACAGCTCCAT 354
Qy      81  LeuGnIleAsnAsnLeuSerGluAspMetAlaThrTyrPheCysAlaArgTrpGlu 100
Db      355 TTGCAGATCAACAACCTCAAAAGTAAGACATGCGCTACATATTTCTGGCAAGAGGAG 414
Qy      101 ValTyrHisGlyTyrValProTyrTrpGlyGlnGlyThrThrValThrValSerSerGly 120
Db      415 GTTACACACGGGTACGTTCTTACTGGGGCCAGAGGACACAGGTACCGGTTCTCTGGC 474
Qy      121 GlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThrGln 140
Db      475 GGTGGCGGTTCTGGGCGGTGGCTCCGCGCGGTGGCTTCGACATCAGCTGACCCAG 534
Qy      141 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSer 160
Db      535 TCTCACAAAATTCCTGTCACCTTCAGTAGAGACAGGGTCAGCATCCTCGCAAGGCGAGT 594
Qy      161 GlnAspValTyrAsnAlaValAlaTrpTyrGlnGlnLysProGlyGlnSerProLysLeu 180
Db      595 CAGGATGATGATAATGCTGTTGCCGTGATACACAGAAACAGAGCAATCTCTAAACTT 654
Qy      181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200
Db      655 CTGATTTACTGGGATCTCCCGGACACTGAGAGCTCTTCCTGCTTCACTGCGAGTGGC 714
Qy      201 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGlnAspLeuAlaValTyr 220
Db      715 TCTGGCGCGGATTTCACTTTACATCAATCAGACAGTGTGAGGCTGAAAGCTGGCAGTTAT 774
Qy      221 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyThrLysLeuGlnIle 240
Db      775 TTCTGTGAGCAACATTTTCGTACTCCATTCACGTTCCGCTGGGGACAAATTTGAGAGTC 834
Qy      241 LysAlaLeuGlnIleSerAsnSerValMetTyrPheSerSerValProValLeuGln 260
Db      835 AAACCTCTAGAGCA----- 848
Qy      261 LysValAsnSerThrThrThrLysProValLeuArgThrProSerProValHisProThr 280
Db      849 -----CATCATCAACCATCACTCAAGA 869
Qy      281 GlyThrSerGlnProGlnArgProGlnAspCysArgProArgGlySerValLysGlyThr 300
Db      870 GGG---CGGCGAGCTTGGCGCGCTGAC---CGCGCACCAAGCGCTGCCACCTGCCGCT 920
Qy      301 GlyLeuAspPheLeuGlnAspProLysLeuCysTyrIleLeuAspGlyIleLeuPheIle 320
Db      921 GGAAGACTTTACCCGCTCATGCGCAGCGCGGCTGGGAAACACTGAGACAGTCCGGCTA 980
Qy      321 TyrGlyValIleIleThrAlaLeuTyrLeuArgAlaLysPheSerArgSerAlaGluThr 340
Db      981 TCCGGTGAGCGGTGGTGGCTCTTACTGCGCGCGCACTGCA----- 1026
Qy      340 ValAlaAsnLeuGlnAspProAsnGlnLeuTyrAsnGlnLeuAsnLeuGlyArgArgGly 360
Db      1027 -----TGGAAACAGGTCAGCA-GGT----- 1046
Qy      360 uGluTyrAspValLeuGlnLysLysArgAlaArgAspProGlnMetCysLysGlnGln 380
Db      1047 -----GATCCGCAACGCCCTGAGCCAGCCCG 1072
Qy      380 ArgArgArgAsnPro-GlnGlnGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaG 400
Db      1073 GCAAGCGCGGAGACTTGGGGAAGCATCGCGAGCAGCCGGAAGCAGCCCGCTGCGCC 1132
Qy      400 LysAlaTyrSerGlnIleGlyThrLysGlyLysArgArgGlyLysGlyHisAsp 418
Db      1133 TGACCTTGCGCGCGCGAGAGCGCTTCTCGCGCAGGCGACCGGCAACAGAC 1188

```

RESULT 14
A57351

```

LOCUS      A57351                2793 bp    DNA    linear    PAT 03-MAR-1998
DEFINITION Sequence 21 from Patent EP0739984.
ACCESSION  A57351
VERSION     A57351.1   GI:3713230
KEYWORDS
SOURCE
  ORGANISM
    1. synthetic construct
      synthetic construct
      other sequences; artificial sequences.
REFERENCE
  1. Wells, W.D., Schmidt, M. and Groner, B.D.
  Bivalent polypeptides containing at least two domains
  Patent: EP 0739984-A 21 30-OCT-1996;
  SAN TUMORFORSCHUNGS GMBH (DE)
  Location/Qualifiers
    1..2793
      /organism="synthetic construct"
      /mol_type="unassigned DNA"
      /db_xref="taxon:32630"
      /clone="PMS238-5-5"
      /note="MOUSE; P. AERUGINOSA; SYNTHETIC"
      1..2679
        /note="unnamed protein product"
        /codon_start=1
        /transl_table=1
        /protein_id="CAA03448.1"
        /db_xref="GI:3713231"
        /translation="MKTALAIATVALAGFATVADYADYKDDDKLHHHHHKLQVOLO
        SGRELYKPRPEYTKISCKASGYFTPTNGAMVAVKQAPQGLKMKGHINTSTGESFPADF
        KGRFDSLSTSAIAYLQINLSEEMATVFCARMEVGVYVWQGGTFTVVSQSG
        GSGGGSGGSGSDIQLTQSHKELSTSGRVSLTCAASQDVNAVAMYOQKQSGSKL
        LIYSASRYTGVPSRFTGSGSPDPFTFTISSVOAEDLAVYFCQHRFTPTFGSGTKL
        ELVALSHHHHHHLEGGSLAALTAHQCHLPLEFTFTRPQRCMEQLQCGYVQRLVA
        LVYLAARLSMNVDQVIRNALASPGSGGDLGEALIREQEPARLALTAAESRRFVROG-
        TGNDEAGANANDEKLOVOLOSGPELKKRGETVTKISCKRSGFTPTNGAMVAVKQAPQ
        GLKMKGHINTSTGESFPADFKGRFDSLSTSAIAYLQINLSEEMATVFCARMEV
        HGEYVWQGGTFTVVSQSGSGGSDIQLTQSHKFLSTSVQEDRVSITCKA
        SQDVNAVAMYOQKQSGSKPLIYSASRYTGVPSRFTGSGSPDPFTFTISSVOAEDL
        AVYFCQHRFTPTFGSGTKLEIKALEDRVPSGAPDASGLALERNPTAEFLGD
        GGSFSTRGTQMTVERLLOAHROLEBERGVFVGHGTEFLPAQSIYAGVARSQD
        LDAIWMGFYIADPDLAVGYADQDEPDARGIRRNALLRVYVPRSLIPFTYSLTILA
        APRAAGEVRLIGHPLPIPLDAITGDEEERGLRLLGMLFARIVVIPSALPTPBRN
        VGDDLDPSISIPDKQKISHLPYASQPKRPREDK"
ORIGIN
Alignment Scores:
Pred. No.:      1,49e-92      Length:      2793
Score:          1308.00      Matches:      275
Percent Similarity: 71.198      Conservative: 24
Best Local Similarity: 65.488      Mismatches: 59
DB:              55.784      Indels:      64
              Gaps:      5
US-09-596-774-7 (1-443) x A57351 (1-2793)
Qy      1  GlnIleGlnLeuValGlnSerGlyProGlnLeuLysLysProGlyGluThrValLysIle 20
Db      115 CAGGTCAACCTGACAGACAGTCTGACCTGAACCTGAAGAACCTGAGAGACAGTCAAGATC 174
Qy      21  SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTrpValLysGlnAla 40
Db      175 TCTCGCAAGGCGCTGGGGTATCTCTTTCACAAACTATGAAATACCTGGGTGAAGAGGCT 234
Qy      41  ProGlyGlnGlyLeuLysTyrPheMetGlyTyrIleAsnThrSerThrGlyGlnSerThrPhe 60
Db      235 CCAAGACAGAGGTTTAAAGTGAATGGGCTGATTAACACTCCACTCGAGAGTCAACATTT 294
Qy      61  AlaAspAspPheLysGlyLysArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80
Db      295 GCTGATGACTTCAAGGAGCGGTTTGAATCTTCTTGGAAACCTGCGCAACACTGCTAT 354
Qy      81  LeuGlnIleAsnAsnLeuSerGluAspMetAlaThrTyrPheCysAlaArgTrpGlu 100
Db      355 TTGCAGATCAACAACCTCAAAAGTAAGACATGCGCTACATATTTCTGTGCAAGATGGAG 414

```

QY 101 ValTyrHisGlyTyrValProTyrTrpGlyGlnGlyThrValThrValSerSergly 120
DB 415 GTTATCCACGGCTACGTTCTTACTGCGGCCCAAGGACCAACGCTTCTCTGCG 474
QY 121 GlyGlyGlySerglyGlyGlyGlySerglyGlyGlySerglyLeuLeuThrGln 140
DB 475 GGTGGCGGTTCTGCTGCGGTGGCTCCGGCGGTGGCTTGCATCATCGAGCTGACCG 534
QY 141 SerHisValPheLeuSerThrSerValGlyAspArgValSerThrCysValValSer 160
DB 535 TCTCAACAATTCCTCTCCACTTCAGTAGAGACAGGCTGATGATCACTCGCAAGGCGAGT 594
QY 161 GlnAspValTyrAsnAlaValAlaTrpTyrGlnGlnLysProGlyGlnSerProLysLeu 180
DB 595 CAGAGTGTATATAGCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 654
QY 181 LeuIleTyrSerAlaSerSerSerglyTyrThrGlyValProSerglyPheThrGlySergly 200
DB 655 CTGATTACTCGGCATCTCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 714
QY 201 SerGlyProAspPheThrPheThrThrIleSerSerValGlnAlaGlyLeuAlaValTyr 220
DB 715 TCTGGCGCGATTCTTCACTTTCACATCAGCAGTGTGCAAGGCTGAGACCTGCGAGTTAT 774
QY 221 PheCysGlnGlnHisPheArgThrPheThrPheGlySerglyThrLysLeuGlnIle 240
DB 775 TTTCTGTCGACACATTTTCTGATCTCATTCACGTTTGGCTGCGGACAAATTTGAGATC 834
QY 241 LysAlaLeuGlnIleSerAsnSerValMetTyrPheSerSerValProValLeuGln 260
DB 835 AAAGCTTAGAGCA----- 848
QY 261 LysValAlaSerThrThrThrLysProValLeuAlaGlnThrProSerProValHisProThr 280
DB 849 ----- 869
QY 281 GlyThrSerGlnProGlnArgProGlnAspCysArgProArgGlySerValLysGlyThr 300
DB 870 GGG---CGGACGCTGCGCGCGCGCTGAC-----CGGACCAACGCGCTGCGACCTGCGCT 920
QY 301 GlyLeuAspPheLeuGlnAspProLysLeuGlyTyrLeuLeuAspGlyIleLeuPheIle 320
DB 921 GGAGACTTTCACCGCTCATCGCCACCGCGCGCGCGGGAACAACATCGAGCACTGCGCTA 980
QY 321 TyrGlyValIle-IleThrAlaLeuTyrLeuArgAlaLysPheSerAlaGluThr 340
DB 981 TCCGGTGCAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1026
QY 340 AlaAlaAlaSerLeuGlnAspProAsnGlnLeuTyrAsnGlnLeuAlaAspGlyArgArgGly 360
DB 1027 -----TGGAAACAGGTCGACCA-GGT----- 1046
QY 360 uGluTyrAspValLeuGlnLysLysArgAlaArgAspProGlnLysGlyLysGlnGly 380
DB 1047 -----GATCGCAACGCCCTGCGCAAGCCCG 1072
QY 380 ArgArgArgArgAspPro-GlnGlnGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaG 400
DB 1073 GCACGCGCGCGACCTGCGCGACAGCATCCGCGACACCGCGAGCAGCGCTGCGCTG 1132
QY 400 LysAlaTyrSerGlnIleGlyThrLysGlyGlnArgArgArgGlyLysGlyHisAsp 418
DB 1133 TGACCTGCGCGCGCGCGCGAGCGAGCGCTTCTGCGCGAGGACCGGCAACGAC 1188

RESULT 15
A57347
LOCUS A57347 3177 bp DNA linear PAT 03-MAR-1998
DEFINITION Sequence 17 from Patent EP0739884.
ACCESSION A57347 GI:3713226
VERSION A57347.1
KEYWORDS
SOURCE synthetic construct

ORGANISM synthetic construct
REFERENCE other sequences; artificial sequences.
AUTHORS 1 Wells, W.D., Schmidt, M. and Groner, B.D.
TITLE Bivalent polypeptides containing at least two domains
JOURNAL Patent: EP 0739884-A 17 30-OCT-1996;
SAN TUMORFORSCHUNGS GMBH (DE)
FEATURES location/Qualifiers
SOURCE 1. 3177
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/clone="PMS240-5-225"
1. 3063
/note="unnamed protein product"
/codon_start=1
/transl_table=1
/protein_id="CA03446.1"
/db_xref="GI:3713227"
/translation="MKKTALAIAYALAGFATVAADYKDDDKLHHHHHKLAVQLOO
SGPELKKRGETVKISCKASGYPTNYGMNWKARPGGLKMGMINSTGSEADPE
KGRDPSLETSANRAYLIQINNLKSEDMATYRCAMWEYHGVYWGCGTLYTVSSGG
GSGGSGGGSDIQLTQSHKFTSVGDRVSTCKASQDVYNAVATQKRGSPKL
LIVASASRYTGVPFRFTGSGGPDFTTISVQAEADLAIVFCQHFPTPTFGSGTGL
EIKALEHHHHHLEGGSLAALTAHQACHLPLETFRHRQPRMGSLAEQCGPVORLVA
LYLAARLSMNQDVDIRNALASPGSGDLGAIIEQEPQARLALTLLAAESERVRQ
TGNDAGAAANDDELQVLOESGPGVQVQSGSLITCTVSGFSLTNYGVHVRQSPCK
GLEMLGTVSGGANDYNTPTFSRLSTINKNSKGVFRPMNSLQSDNTLIYCAALTY
YDFEFAYWQGTIVTVSSGGSGGGSDIQLTQSHKFTSVSGEVSFSCRA
SOSIGTIHWYQDRTPNSPRLIIVYASBSISGIPRSGSGSDFTLSINSVSEDI
ADYVCOONNMPTTFGAGTKLEIKALEHHHHHLEGGSLAALTAHQACHLPLETFR
RQPRMGSLAEQCGPVORLVAALYLAARLSMNQDVDIRNALASPGSGDLGAIIEQ
EQARLALTLLAAESERVRQGTGNDAGAAANDYNTPTFSRLSTINKNSKGVFRPMNS
RNYPTGAFPLEDGGDVSFSTGTQNTWYVERLQHRQLEERGIVFYVGHGFTFLAOS
IVFGVARSDLDIAIWRFYIADPPLAIVADQDEPDARGLRNGLALVYPRSS
LPGFYRSLTLAAEEAGEVERLIGHPLPLDLATGPBEEGRLLETLIGMPLEARTV
VIPSAIPTDPNVDLPDSIIPDKQALISALPDYASQPKRPREDLK"

ORIGIN
Alignment Scores:
Pred. No.: 1.75e-92 Length: 3177
Score: 1308.00 Matches: 275
Percent Similarity: 71.19% Conservative: 24
Best Local Similarity: 65.48% Mismatches: 59
Query Match: 55.78% Indels: 64
DB: Gaps: 5

US-09-596-774-7 (1-443) x A57347 (1-3177)

QY 1 GlnIleGlnLeuValGlnSerGlyProGlnLeuLysLysProGlyGlnThrValLysIle 20
DB 115 CAGGTACAACTGCGACGCTGCGACCTGACCTGAGAGAGCGCTGAGACAGCTCAAGTC 174
QY 21 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTrpValLysGlnAla 40
DB 175 TCTGCAAGGCTCTGGGTATCTTTCAACAACATATGGAATGAATGGAGTGAAGAGCT 234
QY 41 ProGlyGlnGlyLeuLysTrpMetGlyTyrIleAsnThrSerThrGlyGlnSerThrPhe 60
DB 235 CCAGAGACAGGCTTAAAGTGATGGCGCTGATTAACCTTCACCTGCGAGATCAACAATTT 294
QY 61 AlaAspAspPheLysGlyArgArgPheAspSerLeuGlnTrpSerAlaAsnThrAlaLys 80
DB 295 GCTATATGCTTCAAGGAGCGGTGTGACTCTCTTTGGAAACCTTCGCGCAACACCTGCTAT 354
QY LeuGlnIleAsnLeuLysSerGlyAspMetAlaThrTyrPheCysAlaArgTrpGln 100
DB 355 TTGCAAGTCAACAACTTCAAAAGTGAAGACATGCTCATATTTCTGTGCAAGATGGAG 414
QY 101 ValTyrHisGlyTyrValProTyrTrpGlyGlnGlyThrThrValThrValSerSergly 120
DB 415 GTTATCCACGGCTACGTTCTTACTGCGGCGCAAGGACCAACGCTCACCGTTCTCTGCG 474

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 20, 2005, 07:10:49 ; Search time 706.644 Seconds
(without alignments)
3711.128 Million cell updates/sec

Title: US-09-596-774-7

Perfect score: 2345
Sequence: 1 QIQLVQSGPELKKEGETYKI.....LSTATKDTYDALHQTIAPR 443

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool_h/US05956774/runat_20042005_075441_12501/app.query.fasta_1.1230
-DB=N Geneseg -QFMT=fastlap -SUFIX=apr20.rng -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=Dlosum62 -TRANS=human40.cdi
-LIST=45 -LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pro -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US05956774 @CGN_1_1_653 @runat_20042005_075441_12501 -NCPU=6 -ICPU=3
-NO MMAP -LARGQUDERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseg_16Dec04:*

1: geneseqn1980s:*\n2: geneseqn1990s:*\n3: geneseqn2000s:*\n4: geneseqn2001as:*\n5: geneseqn2001bs:*\n6: geneseqn2002as:*\n7: geneseqn2002bs:*\n8: geneseqn2003as:*\n9: geneseqn2003bs:*\n10: geneseqn2003cs:*\n11: geneseqn2003ds:*\n12: geneseqn2004as:*\n13: geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2245	100.0	1479	2	AAT05783 Leader-sc
2	1636.5	69.8	1452	2	AAV73324 Chimeric
3	1501.5	64.0	1284	2	AAV77139 Single ch
4	1501.5	64.0	3058	2	AAV77137 Single ch
5	1312.5	56.0	2070	2	AAT42036 Plasmid p

6	1312.5	56.0	2730	2	AAT42044	Aat42044 Plasmid p
7	1309.5	55.8	2011	2	AAQ28262	Aaq28262 Fv(FRPs)-
8	1308	55.8	1862	2	AAT29413	Aat29413 DGAL4-DT-
9	1308	55.8	1919	2	AAT29412	Aat29412 DGAL4-DT-
10	1308	55.8	2214	2	AAT42038	Aat42038 Plasmid p
11	1308	55.8	2793	2	AAT42039	Aat42039 Plasmid p
12	1308	55.8	2793	2	AAT42043	Aat42043 Plasmid p
13	1308	55.8	3177	2	AAT42041	Aat42041 Plasmid p
14	1305	55.7	1692	2	AAT29409	Aat29409 scFv(FRPs)-
15	1303	55.6	2233	2	AAQ28257	Aaq28257 Fv(FRPs)-
16	1302	55.5	794	2	AAT42034	Aat42034 Plasmid p
17	1302	55.5	2718	2	AAT42042	Aat42042 Plasmid p
18	1302	55.5	2793	2	AAT42040	Aat42040 Plasmid p
19	1285.5	54.8	748	2	AAT05781	Aat05781 ErbB-2-sb
20	1276.5	54.4	748	2	AAQ28256	Aaq28256 Fv heavy
21	1249	53.3	1920	10	ADH34695	Adh34695 CE7-spec1
22	1249	53.3	6834	10	ADH34699	Adh34699 CE7-spec1
23	1015.5	43.3	7654	8	ABX16565	Abx16565 Retrovira
24	1015.5	43.3	7654	10	ABX13168	Abx13168 Retrovira
25	1013	43.2	726	4	AAF30723	Aaf30723 Anti-FIX/
26	1011.5	43.1	1422	2	AAT90509	Aat90509 DNA encod
27	1011.5	43.1	1545	2	AAT90510	Aat90510 DNA encod
28	994	42.4	771	4	AAQ8584	Aaq8584 DNA encod
29	994	42.4	771	9	AAT43523	Aat43523 scFVSA fu
30	981.5	41.9	909	2	AAT36878	Aat36878 741F8 ant
31	981.5	41.9	909	2	AAT91837	Aat91837 Anti-C-er
32	981.5	41.9	909	2	AAT78879	Aat78879 Single ch
33	981.5	41.9	909	2	AAV22598	AAV22598 DNA encod
34	981.5	41.9	909	2	AAV21796	AAV21796 741F8 ant
35	981.5	41.9	909	2	AAV63397	AAV63397 Anti-C-er
36	981.5	41.9	909	10	AAAD61483	AAad61483 741F8 sfv
37	975.5	41.6	777	10	ADG32297	Adg32297 Mouse scf
38	975.5	41.6	987	10	ADG32348	Adg32348 DNA encod
39	972.5	41.5	2001	2	AAQ45088	Aaq45088 Sequence
40	962.5	41.0	1956	2	AAT90512	Aat90512 DNA encod
41	962.5	41.0	2079	2	AAT90513	Aat90513 DNA encod
42	960.5	41.0	777	10	ADG32299	Adg32299 Mouse scf
43	960.5	41.0	987	10	ADG32350	Adg32350 DNA encod
44	930	39.7	1350	2	AAT97844	Aat97844 DNA of a
45	928.5	39.6	800	3	AAA08034	Aaa08034 Antibody

ALIGNMENTS

RESULT 1	
ID AAT05783	standard; CDNA; 1479 BP.
XX	
AC AAT05783;	
XX	
DT 16-MAR-1996	(first entry)
XX	
DE Leader-scFv(FRPs):lyt-2 hinge:zeta CDNA.	
XX	
KW Single chain antibody; scFv; antibody engineering; antitumour;	
KW tumour antigen binding; T-cell receptor; cytotoxic T-lymphocyte;	
KW monoclonal antibody; erbB-2; cancer; cell targeting; gene therapy;	
KW adoptive immunotherapy; ss.	
XX	
OS Synthetic.	
XX	
PH Key	Location/Qualifiers
FT CDS	40..1426
FT	/*tag= a
FT	/product= "Leader-scFv(FRPs):lyt-2 hinge:zeta"
XX	
PN MO9530014-A1.	
XX	
PD 09-NOV-1995.	
XX	
PF 20-APR-1995;	95WO-EP001494.
XX	
PR 02-MAY-1994;	94EP-00810244.

XX (CIBA) CIBA GEIGY AG.

XX Groner B, Moritz D;

XX WPI; 1995-393085/50.

DR P-PSDB; AAR85508.

XX New bifunctional proteins for use in killing tumour cells - contg. a
PT tumour antigen binding domain, a hinge region and a zeta chain derived
PT from a T-cell antigen receptor.

XX Example 1; Page 29-31; 46pp; English.

XX cDNA (AAT05783) codes for a bifunctional protein (AAR85505) consisting of
CC single chain antibody scFv(FRP)5 directed against the tumour erbB-2
CC antigen, a hinge region, and a functional zeta chain obt. from a T-cell
CC receptor. The cDNA is expressed in host cells, esp. cytotoxic T-
CC lymphocytes, providing them with a defined tumour cell specificity
CC enabling targeting to defined tumour cells and MHC-unrestricted and MHC-
CC independent tumour destruction in vitro or in vivo

XX Sequence 1479 BP; 369 A; 395 C; 394 G; 321 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 3e-177 Length: 1479
Score: 2345.00 Matches: 443
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
Gaps: 0

US-09-596-774-7 (1-443) x AAT05783 (1-1479)

QY 1 GlnIleGlnLeuValGlnSerGlyProGluLeuLysPProGlyGlnThrValLysIle 20
DB 94 CAGATCCAGTGGTGACCTGACCTGAGCTGAAGAGCCTGAGAGACAGTCAAGATC 153
QY 21 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrPvalLysGlnAla 40
DB 154 TCCGCGAAGGCGCTCGGATCCTTCCACAACTATGAAATGAATGAGCTGGTAAAGCAAGCT 213
QY 41 ProGlyGlnGlyLeuLysTyrMetGlyTyrIleAsnThrSerThrGlyGlnSerThrPhe 60
DB 214 CCAGAGACGGGCTTAAAGTGGATGGGCTGATTAACCTCCACCTGAGAGATCAACATTT 273
QY 61 AlaAspAspPheLysGlyArgPheAspPheSerLeuGlnThrSerAlaAsnThrAlaTyr 80
DB 274 GCTATGACTTCAAGGAGGAGGCTTGAATCTCTTGGAAACCTGACCAACAGCTGCTAT 333
QY 81 LeuGlnIleAsnAsnLeuLysSerGlyLysAspMetAlaThrTyrPheCysAlaArgTyrGlu 100
DB 334 TTGAGATCAACCACTCAAAAGTGAAGACATGCTCATATTTCTGTGCAAGATGGAG 393
QY 101 ValTyrHisGlyTyrValProTyrTyrGlyGlnGlyThrThrValThrValSerSerGly 120
DB 394 GTTACCAACGGCTACGTTCTTACCTGAGGCCAAGGACACAGGTCACGTTTCTCTGGC 453
QY 121 GlyGlyLysSerGlyGlyGlySerGlyGlyGlyLysSerAspIleGlnLeuThrGln 140
DB 454 GGTGGCGGCTTGTGGGGGTGGCTCCGGCGGTGGCTTGTGCAATCCAGCTGACAG 513
QY 141 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSer 160
DB 514 TCTACAAATTCCTGCTCACTTCACTGAGAGACAGGCTGACATCACCCTGCAAGGCCAGT 573
QY 161 GlnAspValTyrAsnAlaValAlaTyrTyrGlnLysPProGlyGlnSerProLysLeu 180
DB 574 CAGGATGCTGATATGCTGTTGCTGGATCAACAGAAACAGAGCAATCTCCCTAACTT 633
QY 181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200
DB 634 CTGATTACTCGGATCCTCCGGGTACAGTGAATCTCTCTGCTTCACTGGGAGTGGC 693

QY 201 SerGlyProAspPheThrPheThrIleSerSerValGlnIleAspLeuAlaValTyr 220
DB 694 TCTGGCCGGATTCACCTTCAACATCAGAGCTGAGCTGAAGCTGGGAGCTTTAT 753
QY 221 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyThrLysLeuGlnIle 240
DB 754 TTCTGTGAGCAACTTTTGTGATCTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 813
QY 241 LysAlaLeuGlnLysSerAsnSerValMetTyrPheSerSerValProValLeuGln 260
DB 814 AAGCTCTAGAGATCAGCAACTCGGTGATGATCTTCACTTCACTTCACTTCACTTCACT 873
QY 261 LysValAsnSerThrThrThrLysProValLeuArgThrProSerProValHisProThr 280
DB 874 AAGTGAACCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 933
QY 281 GlyThrSerGlnProGlnArgProGluAspCysArgProArgGlySerValLysGlyThr 300
DB 934 GGGACATCTAGCCCGCAGAGACAGAAAGTTGTCGGCCCGCTGCTCAGTGAAGGGGACC 993
QY 301 GlyLeuAspPheLeuGlnAspProLysLeuCysTyrIleLeuAspGlyIleLeuPheIle 320
DB 994 GGATGGACTTCTTCAAGAGATCCCAAACTCTGCTACTGCTAGATGGAATCTTCTATC 1053
QY 321 TyrGlyValIleIleThrAlaLeuTyrIleuArgAlaLysPheSerArgSerAlaGluThr 340
DB 1054 TACGAGTCACTATCAACAGCCCTGATCTGAGAGCAAAATTCAGCAGAGTGCAGAGACT 1113
QY 341 AlaAlaAsnLeuGlnAspProAsnGlnLeuTyrAsnGlnLeuAsnLeuGlyArgArgGlu 360
DB 1114 GCTGCCAAGCTGCAGAGCCCAACCAAGCTCTCAATGAGCTCAATGAGGCGAAGAG 1173
QY 361 GlyTyrAspValLeuGlnLysLysArgAlaArgAspProGluMetGlyGlyLysGlnGln 380
DB 1174 GAATATGACCTCTTGGAGAGAGAGCGGCTCGGATCCAGATGGAGAGCAACAGCAG 1233
QY 381 ArgArgArgAspProGlnGlnGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaGlu 400
DB 1234 AGAGAGAGAAACCCCAAGAGGCGTATTCATGCTCAGCAAGAAAGCAAGATGGCAGAA 1293
QY 401 AlaTyrSerGlnIleGlyThrLysGlyGluArgArgGlyLysGlyHisAspGlyLeu 420
DB 1294 GCTTACAGTGAATCGGACCAAAAGCGAGAGCGGAGGAGGAGGAGGAGGAGGAGCTT 1353
QY 421 TyrGlnGlyLeuSerThrAlaThrLysAspThrTyrAspAlaLeuHisMetGlnThrLeu 440
DB 1354 TACCAGGGTCTCAGCACTGCAACAGAGCAACCTATGATGCCCTGATGAGACCTTG 1413
QY 441 AlaProArg 443
DB 1414 GCCCTTCGC 1422

RESULT 2
AAV73324 standard; DNA; 1452 BP.
ID AAV73324 standard; DNA; 1452 BP.
AC AAV73324;
XX XX
XX 26-FEB-1999 (first entry)
XX XX
DE Chimeric CD19/CTCR DNA.
XX XX
XX Single chain; antibody; B-cell marker; CD19; scFvCD19; gene therapy;
XX XX
XX retroviral vector; tumour-associated antigen; cancer; immunoglobulin;
XX XX
XX CD8 alpha chain; hinge region; theta chain; transmembrane domain;
XX XX
XX intracellular domain; T-cell receptor-CD3 complex; ss.
OS Homo sapiens.
XX XX
XX Synthetic.
PN DE19720152-A1.
XX XX

PD 05-NOV-1998.
 XX 02-MAY-1997; 97DE-01020152.
 XX 02-MAY-1997; 97DE-01020152.
 XX (DELB-) DELBRUECK CENT MOLEKULARE MEDIZIN MAX.
 PI Schroeder A, Schwenkenbecher J, Bargou R;
 XX WPI. 1998-584686/50.
 DR P-PSDB; AAM82345.
 XX
 PT Retroviral vector system for gene therapy - by expression of therapeutic
 XX gene in T cells.
 XX
 PS Claim 9; Page 6; 10pp; German.
 CC This sequence encodes a novel chimeric CD19/CTCR protein which is
 CC composed of an immunoglobulin leader sequence, an anti-CD19 single chain
 CC antibody fragment (SCFvCD19), a hinge region from the CD8 alpha-chain and
 CC the transmembrane and intracellular domain of the T-cell receptor-CD3
 CC complex theta chain. This sequence is used in a retroviral vector system
 CC for gene transfer and expression in primary human T cells which involves
 CC a retroviral vector containing cDNA corresponding to a therapeutic gene,
 CC a promoter selected from the SV40 immediate early promoter or the human
 CC IL-2 or MHC I promoter and a selectable marker, and a packaging cell line
 CC (i.e. a modified retroviral packaging line for enhanced transfer into
 CC primary human T lymphocytes). This vector system in which the cDNA
 CC encodes a single-chain antibody (scFv) specific for a tumour-associated
 CC antigen can be used for gene therapy of cancer, especially by isolating T
 CC cells from the patient's blood, stimulating the T cells in vitro,
 CC introducing the retroviral vector into the T cells and returning the T
 CC cells to the patient, preferably by intravenous injection
 XX
 SQ Sequence 1452 BP; 381 A; 379 C; 391 G; 301 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 7.77e-121 Length: 1452
 Score: 1636.50 Matches: 326
 Percent Similarity: 78.45% Conservative: 38
 Best Local Similarity: 70.26% Mismatches: 79
 Query Match: 69.79% Indels: 21
 DB: 2 Gaps: 5

US-09-596-774-7 (1-443) x AAV73324 (1-1452)

QY 1 GlnIleGlnLeuValGlnSerGlyProGluLeuLysProGlyGluThrValIle 20
 DB 58 GAGGTACAACTGCAAGGAGCTGAGCGTGGTGGAGCGCTCTCACTGAAGATT 117
 QY 21 SerGlyValIleSerGlyThrProPheThrValMetLeuThrValIleGln 40
 DB 118 TCCGTCAGAGGCTTCTGCTATGATCTAGTCTAGTGAATCTGCTGTAAGCAGG 177
 QY 41 ProGlyGlnGlyLeuLeuThrMetGlyTrpIleAsnThrSerThrGlyGluSerThrPhe 60
 DB 178 CTTGGACAGGCTTCTGAGTGGATTGACAGATTGGCTCGAGATGAGTATCTACTAC 237
 QY 61 AlaAspAspPheLysGlyArgPheAspPheSerLeuGlnThrSerAlaAsnThrAlaTyr 80
 DB 238 AATGGAAGACTTCAGAGGTAAAGCCACTGACTGACAGCAATCTCCAGCAGACGCTAC 297
 QY 81 LeuGlnIleAsnAsnLeuLysSerGluAspMetAlaThrTyrPheCysAlaArgTrpGlu 100
 DB 298 ATGCAACTCAGCAGCTTCTGAGCACTCTGCGGCTCTATTCTGTGCAAGACGCGAG 357
 QY 101 Val-----TyrHisGlyTyrValProTyrTrpGlyGlnGlyThrThrVal 115
 DB 358 ACTACGACGCGTGGCCCTTATCTATCTATGACTGCTGCGGCTCAAGAACTCATGTC 417
 QY 116 ThrValSerSerGlyGlyGlySer-----GlyGlyGlyGlySerGlyGlyGly 133

DB 418 ACCGCTCTCCAGCCAAAAGACACCAAGCTTGAAAGGTGAATTTTCAGAACGACGC 477
 QY 134 SerAspIleGlnLeuThrGlnSerHisLysPheLeuSerThrSerValGlyAspArgVal 153
 DB 478 GTAGTATCTTGATCACCACCAATCTCCAGCTTTGGCTGGGCTCTAGGCGAGGCGC 537
 QY 154 SerIleThrCysLysAlaSerGlnAspVal--TyrAsnAla-----ValAlaTrp 169
 DB 538 ACCATCTCTCCAGGAGCCAGCAAGTGTGATTATGATGATGATGATTTATTTGAACCTCG 597
 QY 170 TyrGlnGlnLysProGlyGlnSerProLysLeuLeuIleTyrSerAlaSerSerArgTyr 189
 DB 598 TACCAACAGATTCACAGGACAGCCCAACTCTCATCTATGATGATGATGATGATGATGAT 657
 QY 190 ThrGlyValProSerArgPheThrGlySerGlySerGlyProAspPheThrPheThrIle 209
 DB 658 TCTGGATCCACCCAGGCTTATGAGTGGAGTGGGCTGGAACAGACTTCAACCTCAACATC 717
 QY 210 SerSerValGlnAlaGluAspLeuAlaValTyrPheCysGlnGlnHisPheArgThrPro 229
 DB 718 CATCTGTGAGAAAGGTGATGCTGCAACCTATCATCTGTCAGCAAGTGAAGATCG 777
 QY 230 PheThrPheGlySerGlyThrLysLeuGluIleLysAla----- 242
 DB 778 TGAACGTTCCGTGAGGACCAAGCTCGAGATTCAGAAACAGAACTGATCTCCAGAGAA 837
 QY 243 -----LeuGlnIleSerAsnSerValMetTyrPheSerSerValProValLeu 259
 DB 838 GATCTGAACCTTGAAGATGACAGCACTCGGTATGATCTTCTGTCGCACTCTT 897
 QY 260 GlnLysValAsnSerThrThrThrLysProValLeuArgThrProSerProValHisPro 279
 DB 898 CAGAAAGGAAGCTTACTACTACCAAGCACTGCTGCGAATCTCTCACTGTCGACCT 957
 QY 280 ThrGlyThrSerGlnProGlnArgProGluAspCysArgProArgGlySerValIleGly 299
 DB 958 ACCGGACATCTCAGCCCAAGACCAAGAAATGTCGCCCCGTGCTCAGTGAAGGGG 1017
 QY 300 ThrGlyLeuAspPheLeuGluAspProLysLeuCysTyrLeuLeuAspGlyTleLeuPhe 319
 DB 1018 ACCGATTTGACTTGTCTGAGGATCCCAACTGCTTCTGATGATGATGATGATGATGATGAT 1077
 QY 320 IleTyrGlyValIleIleThrAlaLeuTyrLeuArgAlaLysPheSerArgSerAlaGlu 339
 DB 1078 ATCTACGAGTCATCATCACAGCCCTGATCTGAGACCAAAATTCAGAGGAGTGCAGAG 1137
 QY 360 GlnGluLysTyrAspValLeuGlnLysLysArgAlaArgAspProGluMetGlyGlyGln 379
 DB 1198 GAGGAATATGACGTTCTGAGAAAGAGAGCGGCTGGGATCCAGAGATGGAGGCAACAG 1257
 QY 380 GlnArgArgArgAsnProGlnGlnGlyValTyrAsnAlaLeuGlnLysAspLysMetAla 399
 DB 1258 CAGAGGAGGAGAAACCCCGAGAGGCGTATCAATGCACTGCGCAAGAAAGCAATGCGCA 1317
 QY 400 GluAlaTyrSerGlnIleGlyThrLysGlyGluArgArgArgGlyLysGlyHisAspGly 419
 DB 1318 GAACTTACAGTGAATGCGGACAAAGAGCGAGAGCGGAGAGGCAAGAGGCGCAAGTGGC 1377
 QY 420 LeuTyrGlnGlyLeuSerThrAlaThrLysAspThrTyrAspAlaLeuHisMetGlnThr 439
 DB 1378 CTTTACACAGGCTTCAACATGCGACCAAGACCACTATGATGCTCTGCAATATGACAGCC 1437
 QY 440 LeuAlaProArg 443
 DB 1438 CTGGCCCTCTCC 1449
 RESULT 3
 AAT77139
 ID AAT77139 standard; DNA; 1284 BP.


```

    revised)
    first entry)
    antigen hybrid receptor partial DNA sequence.
    ; single chain antigen; gene therapy; diagnosis
    on; receptor; control region; ss.

-
Location/Qualifiers
1..1276
/*tag= a
/notes= "partial hybrid receptor"
1..12
/*tag= b
/notes= "partial leader sequence"
13..1285
/*tag= c
/product= "hybrid.receptor"
/notes= "single chain antigen hybrid receptor"

```

Alignment Scores:

Pred. No.:	3, 77e-110	Length:	1284
Score:	1501.50	Matches:	297
Percent Similarity:	76.29%	Conservative:	44
Best Local Similarity:	66.44%	Mismatches:	79
Query Match:	64.03%	Indels:	27
DB:	2	Gaps:	6

US-09-596-774-7 (1-443) X AAT77139 (1-1284)

QY	1	GlnIleGlnLeuValGlnSerGlyProGlnIleuLeuLysProGlyGlnIleuValIle	20
Db	13	CAGGCTTCAACTGCACAGACTGGGGCTGGCTTGAGAGCCTGGGGCTTCAGTGAAGCTG	72
QY	21	SerCysLysAlaSerGlyTyrProMetThrAsnTyrGlyMetAsnTyrValLysGlnAla	40
Db	73	TCCTCAAGGCTTCTGGCTACACCTTCCACAGCTACTGGATGCACTGGGTGAAGCAGAG	132
QY	41	ProGlyGlnGlyLeuLysTyrMetCylTyrPheAsnThrSerThrGlyGlnSerThrPhe	60
Db	133	CTTGAGCAGAGCCTTGAGTGGATTGGAGATTGATCTCAATATAGTGCGTACTGAATGC	192
QY	61	AlaAspAspPheLysGlyArgPheAspPheSerLeuGlnThrSerAlaAsnThrAlaTyr	80
Db	193	AATGAGAACTTCAGAGCAGAGCCACACTGACTGTACGACAAACCTCCAGCAGACGCTTC	252
QY	81	LeuGlnIleAsnAsnLeuLysSerGlyAspMetAlaThrTyrPheCysAlaArgTyrGln	100
Db	253	ATGCAAGCTCACAGCAGCTGCATCTGAGAGACTCTGGGCTATATATATGGCAAGATACGAT	312
QY	101	ValTyr---HisGlyTyrValProTyrTyrGlyGlnGlyThrThrValThrValSerSer	119
Db	313	TACTACGGTAGTACTACTTGTACTACTCGGGCCAGAGGACACACGCTCACCTTCCTCA	372
QY	120	GlyGlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThr	139
Db	373	GGTGAAGCGGCTTCAGGCGAGAGTGGCTCTGGCGGTGCGGATCTCAGGCTTGGGACA	432
QY	140	GlnSerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLysAla	159
Db	433	CAG---GAATCTGCATCACCACATCACCCTGGTGAACAGTACACTCATCTGTGGCTCA	489
QY	160	SerGlnAspValTyr-----AsnAlaValAlaTyrTyrGlnGlnLysProGlyGln	176
Db	490	AGTACTGGGGCTGTACAACATAGTATACCAACATGGGTGCCAAGAAAACCATGATCT	549
QY	177	SerTyrSerLeuLeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPhe	196
Db	550	TTATTCATCGTCTCAATAGGTGTACCAACAACAGAGCTCAGAGGTCTCCGCAATTTC	609
QY	197	ThrGlySerGlySerGlyProAspPheThrPheThrIleSerSerValGlnAlaGlnAsp	216
Db	610	TCAGGCTCCCGATGGATGGAGCAAGGCTGCCCTCACCATCACAGGGGACACAGCTAGGAT	669
QY	217	LeuAlaValTyrPheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyThr	236
Db	670	GAGCAATATATTTCTGTGCTCTATGTGACGACCAATTGGGTGGTGGGAGAGAAC	729
QY	237	LysLeuGlnIleLysAlaLeuGlnIleSerAsnSerValMetTyrPheSerSerValVal	256
Db	730	AAACTGACTGTC-----CTCGAG-----	747
QY	257	ProValLeuGlnLysValAsnSerThrThrThrLysProValLeuArgThrProSerPro	276
Db	748	-----GAATTCACACTCTCCAAAGCAGACGCTGCCGAATCCCTCACCT	789
QY	277	ValHisProThrGlyThrSerGlnProGlyLanArgProGlyAspCysArgProArgGlySer	296
Db	790	GTGCACCTTACCGGAGCATCTCACGCCCCAGACGACAGAAATGTGGCCCCGGTGGCTCA	849
QY	297	ValLysGlyThrGlyLeuAspPheLeuGlnAspProLysLeuCysTyrLeuLeuAspGly	316
Db	850	GTGAAGGGAGACCGGATTGGACTTC-----GACCCAAACTCTGCTACTTGTGATGTA	903
QY	317	IleLeuPheIleTyrGlyValIleIleThrAlaLeuTyrIleuArgAlaLysPheSerArg	336
Db	904	ATCTCTTTCATCTACGAGATCATATACAGCCCTGTACCTGAGACGAAATTCAGCAGG	963
QY	337	SerIleGlnThrAlaAlaAsnLeuGlnAspProAsnGlnLeuTyrAsnGlyLeuAsnLeu	356
Db	964	AGTGCAGAGACTGCTGCCAATCTGCAGAGCCCAACACAGCTCTCAATAGACTCAATCTTA	1022
QY	357	GlyArgArgGlnGlnTyrAspValLeuGlnLysLysArgAlaArgAspProGlnMetCyl	376


```

Db      1024 GGGGGAAGAGGATATGACGCTCTTGGAAGAAAGCGGCTCCGGATCCAGAGATGGGA 1083
Qy      377 GLYISGLINGINAAGARGASAPProGInGUgLYValTYRAsNALeUGInLYAsP 396
Db      1084 GGGAAACGCGAGAGAGAGAACCCCGAGAGGCAATATGCACTGCAGAAAGAC 1143
Qy      397 LysERAAGIUAATYRserGluILEGlyThrLYSGIUAARGARGGLYLYSGIY 416
Db      1144 AACATGGCAGAGCCTACAGTGAATCGGCACAAAAGCGAGAGCGGAGAGGAGGG 1203
Qy      417 HisAPGlyLeuTYRGLNGlyLeuSerThrAlaThrLYsAsPThrTYRAsPALeUHis 436
Db      1204 CAGATGGCCTTTACAGAGGCTCAAGCACTGCACCAAGAGACACTATGATGCCCTGCAT 1263
Qy      437 MetGInThrLeuAlaProArg 443
Db      1264 ATGCAGACCCCTGGCCCTCGC 1284

RESULT 4
AAT77137
ID AAT77137 standard; DNA; 3058 BP.
XX
AC AAT77137;
XX
DT 25-MAR-2003 (revised)
DT 04-MAR-1998 (first entry)
XX
DE Single chain antigen hybrid receptor DNA.
XX
KM Hybrid receptor; single chain antigen; gene therapy; diagnosis;
KM signal conduction; receptor; control region; ss.
XX
OS Synthetic.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT misc_signal /tag= a
FT /note= "Noted in specification"
FT CDS /tag= b
FT /note= "This sequence contains an intron"
FT sig_peptide /tag= d
FT /label= "leader sequence"
FT /note= "This sequence contains an intron"
FT intron /tag= c
FT /tag= c
FT mat_peptide /tag= e
FT /product= "single_chain_antigen_hybrid_receptor"
FT poly_a_site /tag= f
XX
XX NO9720938-A2.
XX
XX 12-JUN-1997.
XX
XX 03-DEC-1996; 96MO-DB002334.
XX
XX 05-DEC-1995; 95DE-01045351.
XX
XX (UYFR-) UNIV FREIBURG KLINIKUM ALBERT-LUDWIGS.
XX
XX Mertelmann R, Kulmburg P, Rosenthal F;
XX
XX MPI, 1997-319784/29.
XX
XX P-PSDB; AAW24025.
XX
XX Cells with hybrid receptor having extracellular and intracellular regions
XX of different origins - useful in gene therapy and diagnosis of tumours.
XX

```

```

PS      Example 3; Fig 4; 46pp; German.
XX
CC      This DNA sequence encodes a novel single chain antigen hybrid receptor
CC (HR) and contains an extracellular domain specific for the hapten 4-
CC hydroxy-5-Iodo-3-nitrophenyl acetate (NIP) coupled to the C8- $\alpha$ 
CC region and the transmembrane and signal-conducting intracellular parts of
CC the C8- $\alpha$  molecule. Such hybrid receptors comprise a receptor part
CC localised on the outside of the cell and specific to a particular signal
CC molecule and a receptor part originating from another receptor, localised
CC on the inside of the cell and capable of setting off a signal inside the
CC cell. The cell should also contain at least one other gene construct with
CC a control region which can interact with the signal sent out by the
CC hybrid receptor and thereby control expression of a transgene bound to
CC this control region. Such cells are useful in gene therapy or for
CC diagnostic purposes. (Updated on 25-MAR-2003 to correct PR field.)
XX
SQ      Sequence 3058 BP; 689 A; 831 C; 853 G; 685 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:      1,08e-109      Length:      3058
Score:          1501.50      Matches:      297
Percent Similarity: 76.29%      Conservative: 44
Best Local Similarity: 66.44%      Mismatches: 79
Query Match:      64.03%      Indels:      27
DB:              2      Gaps:      6

US-09-596-774-7 (1-443) x AAT77137 (1-3058)
Qy      1 GlnILEGInLeuValGInSerGlyProGInLeuLYsPProGlyGInThrValLYsILE 20
Db      862 CAGGTCCAACTGCAGCAGCTGTGGGCTGAGCTTGAAAGCTGGGCTTCAGTAAGCTG 921
Qy      21 SerCYsLYsAlAsSerGlyTYRProPheThrAsnTYRGLYMetAsnTPValLYsGlnAla 40
Db      922 TCCGCAAGGCTCTTGCTACACCTTACCACTGATCGATGATCGATGATGATGATGATG 981
Qy      41 ProGlyGInGlyLeuLYsTPProMetGlyTYRILEAsnThrSerThrGlyLeuSerThrPhe 60
Db      982 CCTGGACGAGCGCTTGAAGTGAATGATGATGATGATGATGATGATGATGATGATGATG 1041
Qy      61 AlaAspAspPheLYsGlyARGPheAspPheSerLeuGInThrSerAlaAsnThrAlaTYR 80
Db      1042 AATGAGAAAGTTCAAGACAGAGGCCACACTGACTGTAGCAAAACCTCCAGACAGCTTAC 1101
Qy      81 LeuGInILEAsnLeuLYsSerGInAspMetAlaThrTYRPhCYsAlaArgTPGIn 100
Db      1102 ATGCAGCTCCAGCAGCTGACATCTGAGACCTCGGCTATATTTGTGCAAGATACGAT 1161
Qy      101 ValTYR---HisGlyTYRValAlProTYRTPGInGlyThrThrValThrValSerSer 119
Db      1162 TACTACGGTACGTACTCTTGACTTACTGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGG 1221
Qy      120 GlyLYsLYsSerGlyGlyGlyGlySerGlyGlyGlyGlySerAspILEGInLeuThr 139
Db      1222 GGTGAGGCGCGTTCAAGCGGAGGAGTGGCTGCGGCTGCGGATCTCAGAGCTTGGAGCA 1281
Qy      140 GInSerHisLYsPheLeuSerThrSerValGlyAspArgValSerILEThrCYsLYsAla 159
Db      1282 CAG---GAATCTGCATCACCACATCTGAGTGAACAGTCAACATCTGTCGCTCA 1338
Qy      160 SerGInAspValTYR-----AsnAlaValAlaAlaTrpTYRGLNGInLYsPProGlyGln 176
Db      1339 AGTACTGGGGCTGTACACTAGTACTAGTACCACTGGGTCCAAAGAAAACAGATCAT 1398
Qy      177 SerProLYsLeuLeuILETYRserAlAserSerArgTYRThrGlyValAlProSerArgPhe 196
Db      1399 TTATTCAGTGTCTAATATAGGTGATCAACCAACGAGCTCCAGGTGTTCCGACAGATT 1458
Qy      197 ThrGlySerGlySerGlyProAspPheThrPheThrILESerSerValGlnAlaGlyAsp 216
Db      1459 TCAGGCTCCCTGATGTAGAGCAAGAGCTGCCCTCACCTACAGGAGGAGGAGGAGGAGGAGT 1518
Qy      217 LeuAlaValTYRPhCYsGInGInILEPheArgThrProPheThrPheGlySerGlyThr 236

```

```

Db 1519 GAGGCAATATATTCTGTCTATGATGACAGCAACCATGAGTGTGCGTGAGAGAACCC 1578
Qy 237 LysLeuGluIleIysAlaLeuGluIleSerAsnSerValMetYrPheSerSerValVal 256
Db 1579 AATCTGACTGTC-----CTCAG----- 1596
Qy 257 ProValIleuGlnIysValAsnSerThrThrIysProValIleuArgThrProSerPro 276
Db 1597 -----GAATTCACCTACTACCAAGCCAGTCTCGCAACTCCCTCACT 1638
Qy 277 ValHisProThrGlyThrSerGlnProGlnArgProGluAspCysArgProArgGlySer 296
Db 1639 GTGCACCTTACCTCGGACATCTCAGCCCAAGACCAAGATTGTCGGCCCGTGCTCA 1698
Qy 297 ValIysGlyThrIleuAspPheLeuGluAspProIysLeuCysYrIleuLeuAspGly 316
Db 1699 GTGAAGGGAGCGGATTTGAGCTTC-----GACCCCAAACTCTGCTACTTGTAGATGA 1752
Qy 317 IleuPheIleYrGlyValIleIleThrAlaLeuYrIleuArgAlaIysPheSerArg 336
Db 1753 ATCTCTTTCATCTACGAGATCATCATCAGCCCTGTACTGAGACCAAAATTCAGCAGG 1812
Qy 337 SerAlaGluTrpAlaAlaAsnLeuGlnAspProAsnGlnLeuYrAsnGluLeuAsnLeu 356
Db 1813 AGTCAGAGACTGCTGCCAACCTCAGAGACCCCAACAGCTCTCAATGACTCAATCTA 1872
Qy 357 GlyArgArgGluGluYrAspValIleuGlnIysIysAlaArgAspProGluMetGly 376
Db 1873 GGGGGAAGAGAGAAATATGACTGTTTGAGAAAGAGCGGCTCGGATCCAGAAATGAG 1932
Qy 377 GlyIysGlnIleuArgArgAsnProGlnIleuGlyValYrAsnAlaIleuGlnIysAsp 396
Db 1933 GGCAAAACGACGAGAGAGAGAACCCCAAGAGCATATGCATGACCTGCAAGAAAGAC 1992
Qy 397 LysMetAlaGluAlaYrSerGluIleGlyThrIysGlyIleuArgArgGlyIysGly 416
Db 1993 AACATGGCAGAAAGCTACAGTACAGTACGAGCAAAAGCGAGCGAGCGCAAGGGG 2052
Qy 417 HisAspGlyLeuYrGlnGlyLeuSerThrAlaThrIysAspThrYrAspAlaLeuHis 436
Db 2053 CACGATGGCTTTTCCAGGAGCTCAGCATGCCACCAAGACCACTTGTATGCTTCAT 2112
Qy 437 MetGlnThrIleuAlaProArg 443
Db 2113 ATCGACGACCTGCGCCCTCGC 2133

RESULT 5
AAT42036
ID AAT42036 standard; DNA; 2070 BP.
XX
AC AAT42036;
XX
DT 29-JAN-1997 (first entry)
XX
DE Plasmid pSW202-5 fragment encoding scFv(FRP5)-ETA.
XX
KM Single chain antibody; scFv; monoclonal antibody; Mab; EGF; erbB-2;
KM epidermal growth factor; receptor; plasmid pSW202-5; cancer; therapy;
XX
XX antitumour; exotoxin A; ETA; ss.
OS Mus; sp.
OS Pseudomonas; aeruginosa.
OS Synthetic.
OS Chimeric.
XX
XX
FH Key Location/Qualifiers
FT CDS 1..1956
FT FT /*tag= a
FT sig_peptide 1..63
FT FT /*tag= b
FT FT /note= "Ompa signal peptide"
FT mat_peptide 64..1953

```

```

FT EP739984-A1. /*tag= c
FT /product= "scFv(FRP5)-ETA"
XX
XX
XX 30-OCT-1996.
XX
XX 26-APR-1995; 95SEP-00106275.
XX
XX 26-APR-1995; 95SEP-00106275.
XX
XX (SANT-) SAN TUMORFORSCHUNGS GMBH.
XX
XX Wels W, Schmidt M, Groner B;
XX
XX WPI; 1996-478748/48.
XX
XX P-PSDB; AAM05156.
XX
XX Bivalent fusion proteins that bind epidermal growth factor receptor or
XX analogues - and comprise at least two different cell surface binding
XX domain(s), useful for tumour therapy.
XX
XX Example 7, Page 21-22; 52pp; English.
XX
XX A fragment (AAT42036) of plasmid pSW202-5 codes for scFv(FRP5)-ETA
XX (AAM05136), comprising the single-chain binding region of murine
XX monoclonal antibody FRP5, which is specific for human epidermal growth
XX factor receptor erbB-2, joined to exotoxin A (ETA). A fragment of plasmid
XX pMW152-5 (see also AAT42034) encoding scFv(FRP5) (AAM05134) was ligated
XX to pSW200, a plasmid contg. the Pseudomonas aeruginosa PAK ETA gene, to
XX produce pSW202-5. This can be utilised in the construction of bacterial
XX expression vectors (see also AAT42037-44) encoding bivalent fusion
XX proteins (AAM05137-44) useful as antitumour agents
XX
XX Sequence 2070 BP; 402 A; 675 C; 628 G; 365 T; 0 U; 0 Other;
XX
XX Alignment Scores:
XX Pred. No.: 7,58e-95 Length: 2070
XX Score: 1312,50 Matches: 284
XX Percent Similarity: 68,59% Conservativity: 13
XX Best Local Similarity: 65,59% Mismatches: 55
XX Query Match: 55,97% Indels: 83
XX DB: 2 Gaps: 8
XX
XX US-09-596-774-7 (1-443) x AAT42036 (1-2070)
Qy 1 GlnIleGlnLeuValGlnSerGlyProGluLeuIysIysProGlyValThrValIysIle 20
Db 115 CAGGTACAACTGACGACAGTCTGACCTGAACCTGAAGAACCTGGAGAGACAGTCAAGATC 174
Qy 21 SerCysIysAlaSerGlyYrProPheThrAsnYrGlyMetAsnTrpValIysGlnAla 40
Db 175 TCCTGCAAGGCTCTGGGTRTCTTTCAAACTATGATGAATGAATCACTGGTGAAGCAGCT 234
Qy 41 ProGlyGlnGlyLeuIysTrpMetGlyTrpIleAsnThrSerThrGlyIleSerThrPhe 60
Db 235 CCAGGACAGGGTTTAAAGTGAATGGGCTGATTAACACTCACTGAGAGCAACATTT 294
Qy 61 AlaAspAspPheIysGlyIysArgPheAspPheSerIleuGluThrSerAlaAsnThrIaYr 80
Db 295 GCTGATGACTTCAGAGGACGGTTTGACTTCTCTTGGAAACCTGCGCAACCTGCTAT 354
Qy 81 LeuGlnIleAsnAsnLeuIysSerGluAspMetAlaThrYrPheCysAlaArgTrpGlu 100
Db 355 TTGCAGATCAACAACCTCAAAAGTGAACAGCTCAATATTTCTGTCCAAAGATGGAG 414
Qy 101 ValYrHisGlyYrValProTrpTrpGlyGlnGlyIleThrValThrValSerSerGly 120
Db 415 GTTTACCAAGGCTACGTTCTTACTGAGGCGCAAGGAGCAACGCGTACCGTTCTCTGGC 474
Qy 121 GlyIysIysSerGlyGlyGlyIysSerGlyIysGlyIysSerAspIleGlnLeuThrGln 140
Db 475 GGTGGGGTTCTGTGGCGGTGGCTCCGCGGTGGCGGTTCGACATCCAGCTGACCCAG 534

```

```

QY      141 SerHisysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLeuAlaSer 160
DB      535 TCTCAAAATTCCTGCTCACTTCAGTAGAGACAGGCTCAGATCACTGCAAGCCGAT 594
QY      161 GlnAspValTyrAsnAlaValAlaIlePTrgIngluLysProGlyGlnSerProLysLeu 180
DB      595 CAGATGATGATATGCTGTTGCTCGTATCAACAGAAACAGACATCTCTAACTT 654
QY      181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerIly 200
DB      655 CTGATTTACTCGGATCTCCCGGTACACTGAGTCCCTTCCTGCTTCACTGCGAGTGC 714
QY      201 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 220
DB      715 TCTGGGCGGATTTCACTTTTCACTTCACATCGAGTGTGCGAGCTGAAAGACTGGCACTTAT 774
QY      221 PheCysGlnGlnAlaPheArgThrProPheThrPheGlySerGlyThrLysLeuGluIle 240
DB      775 TTTCTGTAGCAACATTTTCTGATCTCATTCAGTTCCGCTCGGGGCAAAATTGGAGATC 834
QY      241 LysAlaLeuGluIleSerAsnSerValMetTyrPheSerSerValValProValLeuGln 260
DB      835 AAAGCTCTAGACA----- 848
QY      261 LysValAsnSerThrThrThrLysProValLeuArgThrProSerProValHisProThr 280
DB      849 -----CATTCACTACCACTCACTGAGA 869
QY      281 GlyThrSerGlnProGlnArgProGluAspCysArgProArgGlySerValLysGlyThr 300
DB      870 GGG---GGGACCGTGGCGCGCGTGAC-----GGCGACACAGCGCTGCCACCGCGCT 920
QY      301 GlyLeuAspPheLeuGluAspProLysLeuGlyTyrLeuLeuAspGlyTyrLeuPheIle 320
DB      921 GGAGACTTTTCACTCCGTCATCGCCAGCCGCGCGCTGGGAAACACTGAGACATGCGGCTA 980
QY      321 TyrGlyValIle-IleThrAlaLeuTyrTyrLeuArgAlaLysPheSer----- 335
DB      981 TCCGCTGACAGCGGCTGCTGCTGCTTACTTACCTGCGCGGACTGTCTATGAAACACAGTCCA 1040
QY      336 -----ArgSerAlaGluThrAla-----AlaAs 343
DB      1041 CCAGGTATCGGCAACCGCCCTGGCGACGCCGCGGCGGCGACCTGGGCGAAGCAT 1100
QY      343 nLeuGlnAspProAsnGlnLeuTyrAsnGluLeuAsnLeuGlyArgArgGluTyrAs 363
DB      1101 CCGCGAGCAGCGGAGCAGCGCCGCTGCTGACCTT-GGCCGCGCGCGAGAG----- 1154
QY      363 pValLeuGluLysLysAspArgAlaArgAspProGluMetGlyGlyLysGlnGlnArgArgAr 383
DB      1155 -----CGAGCGCTTCTGTCG-----CGAGGCGACCGGCAACGACGA-- 1190
QY      383 GAsnProGlnGluGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaGluAlaTyrSe 403
DB      1190 ----- 1190
QY      403 rGluIleGlyThrLysGlyGluArgArgGlyLys 415
DB      1191 -----GGCCGCGCGCGCAACCGCCGACGTGTGAG 1220

```

RESULT 6

ID AAT42044 standard; DNA; 2730 BP.

AC AAT42044;

DT 29-JAN-1997 (first entry)

XX Plasmid pMS246-5-5 encoding scFv2 (FRP5/FRP5)-ETA (version 3).

XX Single chain antibody; scFv; monoclonal antibody; Mab; EGF; erbB-2;

KW epidermal growth factor; receptor; plasmid pMS246-5-5; cancer;

```

KW      exotoxin A; ETA; antitumour; ss.
XX      Mus: gp.
OS      Pseudomonas; aeruginosa.
OS      Synthetic.
OS      Chimeric.
XX      Key
FH      CDS
FT      Location/Qualifiers
FT      1..2700
FT      sig_peptide
FT      1..63
FT      /tag= a
FT      /note= "Ompa signal peptide"
FT      mat_peptide
FT      64..2697
FT      /tag= c
FT      /product= "scFv2 (FRP5/FRP5)-ETA"
XX      EP739984-A1.
XX      30-OCT-1996.
PD      26-APR-1995; 95EP-00106275.
XX      26-APR-1995; 95EP-00106275.
XX      26-APR-1995; 95EP-00106275.
XX      (SANT-) SAN TUMORFORSCHUNGS GMBH.
XX      Wels W, Schmidt M, Groner B;
XX      WP1; 1996-478748/48.
XX      DR P-PSDB; AAM05144.
XX      Bivalent fusion proteins that bind epidermal growth factor receptor or
PT      analogues - and comprise at least two different cell surface binding
PT      domain(s), useful for tumour therapy.
XX      Example 12; Page 43-45; 52pp; English.
XX      A fragment (AAT42044) of bacterial expression plasmid pMS246-5-5 codes
CC      for scFv2 (FRP5/225)-ETA (AAM05144), and is obtd. by ligating 2 copies of
CC      DNA coding for the single-chain binding region of murine monoclonal
CC      antibody FRP5 (see also AAT42034) and portions of Pseudomonas aeruginosa
CC      exotoxin A. pMS246-5-5 can be utilised in the prodn. of bivalent fusion
CC      protein in bacterial (esp. E. coli) host cells. Such fusion proteins (see
CC      also AAM05138-43) bind the epidermal growth factor receptor and are
CC      useful as antitumour agents
XX      SQ
SQ      Sequence 2730 BP; 567 A; 832 C; 798 G; 533 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1,06e-94 Length: 2730
Score: 1312.50 Matches: 284
Percent Similarity: 68.59% Conservative: 13
Best Local Similarity: 65.59% Mismatches: 55
Query Match: 55.97% Indels: 83
DB: 2 Gaps: 8
US-09-596-774-7 (1-443) x AAT42044 (1-2730)
QY      1 GlnIleGlnLeuValGlnSerGlyProGluLeuLysLysProGlyGluThrValLysIle 20
DB      115 CAGGTACACAGCGACGAGTGTGACCTGAACTGAAAGAGCTGAGAGACGTCAAGATC 174
QY      21 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValLysGlnAla 40
DB      175 TCTGTCAAGGCTCTGGGTATCTTTCACAAACTATGAACTGAGCTGTAAGACAGCT 234
QY      41 ProGlyGlnGlyLeuLysTyrPheMetGlyTyrIleAsnThrSerThrGlyLysSerThrPhe 60
DB      235 CCAGGACAGGGGTTTAAAGTGAATGGCTGATTAACCTTCACTGAGAGATCAACATTT 294
QY      61 AlaAspAspPheLysGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80

```


Sequence 2011 BP; 383 A; 653 C; 620 G; 355 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.27e-94	Length:	2011
------------	----------	---------	------

Score:	1309.50	Matches:	280
--------	---------	----------	-----

Percent Similarity:	68.36%	Conservative:	16
---------------------	--------	---------------	----

Best Local Similarity:	64.67%	Mismatches:	48
Custom Watch:	57.04%	Match:	51

Query Match:	55.84%	Indels:	91
Query Match:	55.84%	Indels:	91

DB: 2
Geprüf: 8

US-09-596-774-7 (1-443) x AAQ28262 (1-2011)

QY	1	GlnIleGlnLeuValGlnIleSerGlyProGluLeuValLysProGlyValThrValLysIle	20
Db	97	CAGATCAACCTGCAGCAGTGTGGACCTGAACCTGAAGAAAGCTGGAGAACAGTCAAGATC	156
QY	21	SeSCysAlaAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValLysGlnAla	40
Db	157	TCTTGCAGAGGCTCTGGGATTCCTTTCAAACTATGGAAATGAAGTGGGTGAAACAGGCT	216
QY	41	ProGlyGlnGlyLeuLysTyrPheGlyTyrPileAsnThrSerThrGlyLysSerThrPhe	60
Db	217	CCAGGACAGGGTTTAAAGTGAATGGGCTGGATTAAACACTCAGCTGAGAGACTCAACATTT	276
QY	61	AlaAspAspPheLeuGlyIvalrPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr	80
Db	277	GCTGAATGACTTCAGAGGACGGTTGACTTCTCTTTGAAACCTCTGCCAACAGCTCTAT	338
QY	81	LeuGlnIleAsnAsnLeuLysSerGluAspMetAlaThrTyrPheCysAlaArgTyrGlu	100
Db	337	TTGCACATGAACAACCTCBAAGTGAAGACATGGCTACATATTTCTGTGCAGAAATGGAG	398
QY	101	ValTyrHisGlyTyrValProTyrTyrGlyGlnGlyIleThrValThrValSerSerGly	120
Db	397	GTTTACCAAGGCTACCTTCCTTACTAGGGGCCAAGGAGACAGGTCACCGTTTCTCTGGC	456
QY	121	GlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlyGlySerAspIleGlnLeuThrGln	140
Db	457	GGTGGGGGTTCTGGTGGCGGTGGCTCCGGCGGTGGCGGTCTGCACATCAGCTTACCCAG	516
QY	141	SerHisLeuPheLeuSerThrSerValGlyAspArgValSerIleThrCysAlaAlaSer	160
Db	517	TCTCAAAATTCCTGTCCACTTCAGTAGGAGCAGGGTCAGATCAGCTCGAAGCCAGT	576
QY	161	GlnAspValTyrAsnAlaValAlaTyrTyrGlnGlnLysProGlyGlnSerProLysLeu	180
Db	577	CAGGATGTGATATATCTGTGCTGGTGTATCAAGAAACAGACAGATCTCTAAACTT	636
QY	181	LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly	200
Db	637	CTGATTATTCCTGGCAGTCCCTCCCGGTACACTGGAGTCCCTTCTCGCTTCACTGGCAGTGGC	696
QY	201	SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr	220
Db	697	TCTGGGCGCGAATTTCACTTTCACATCGACAGATGGCAGGCTGAAGACCTGGCAGTTAT	756
QY	221	PheCysGlnGlnHisIlePheArgThrProPheThrPheGlySerGlyThrLysLeuGlnIle	240
Db	757	TTCTGTCAACACATTTTGTACTCATTCACAGTTCCGCTGGGGACAAATATGGAGATC	816
QY	241	LysAlaLeuGlnIleSerAsnSerValMetTyrPheSerSerValProValLeuGln	260
Db	817	AAAGCTCTAGAG-----828-----	
QY	261	LysValAsnSerThrThrThrLysProValLeuArgThrProSerProValHisProThr	280
Db	829	-----GGCGGACGCTGGCGCGCGCTGAC-CGCGACACAGGCTGCCACTCCGCGCT	878
QY	281	GlyThrSerGlnProGlnArgProGluAspCysArgProArgGlySerValLysGlyThr	300
Db	879	GGAGACTTTCACCCGTCATCGCCA-----GCCGCGCGGCTGGGAACA	920
QY	301	GlyLeuAspPheLeuGlnAspProLysLeuCysTyrIleLeuLeuAspGlyIleLeuPheIle	320

D	b		921	-	-----ACTGAGCAGTGGCTA	938
O	y		321	TyrGIyValle-IleThrAlaleuTYrLeuArGaLaLysPheSer-	-----	335
D	b		939	TCCGTGCACGGCGTGTCGCCCTCTACCTGCGCGCGACATGACGAACCAAGTCA	998	
O	y		336	-ArgSerIaGlunHraha-----	AlaAs	343
D	b		999	CNAGTGTATCCGAAACGCTGGCCACCCCCGACGCGCGCGAACCTGGCGAAGCAT	1058	
O	y		343	nLeuGlnAspProAbnGlnLeuTYrAenGluLeuAenLeuGlyARgRgluGluTYrAs	363	
D	b		1059	CCGGAGACACCCGAGAGAGCCCGCTCTGGCCCTGACCCT-GGCCGCGCGGAG--	1112	
O	y		363	pValleuGluLylalysArGaLaArGAspProGluMetGlyLYlYlsGlnGlnArGAR	383	
D	b		1113	-CGAGCGCTTCGTCCG-----GCAGGGCACCGGCACGACGA--	1148	
O	y		383	GAsnProGlnGluGlyValTYrAsnAlalaLeuGlnLysAspLysMetAlaGluAlaTYrSe	403	
D	b		1148	-	-----	1148
O	y		403	RgluileGlyThrylsGlyGluArGARgLyLYe	415	
D	b		1149	-GCGCGCGCGGCCAACGCCAGCTGTGTGAG	1178	
 RESULT 8 AAT29413 ID AAT29413 standard; DNA, 1862 BP.						
A	C		AAT29413;			
X	X		16-OCT-2003 (revised)			
D	T		19-AUG-1996 (first entry)			
X	X					
D	E		DGAL4-DT-scfv(FRP5) gene in pSW55-GD5.			
X	X					
N	ucleic acid transfer system; gene transfer; gene therapy; cell targeting; multidomain protein; vector; cancer; GAL4; diphtheria toxin; single chain antibody; scFv; ss.					
S	accharomycetes; cerivisiae. Corynebacterium; diphtheriae. Mus sp. OS Chimeric. Synthetic.					
F	H Key		Location/Qualifiers			
P	T CDS		1..1908			
P	T FT		/tag= a			
P	T FT		/product= "GAL4-DT-scfv(FRP5) fusion protein"			
X	X					
M	OS613599-Al.					
X	X					
P	D 09-MAY-1996.					
X	X					
P	F 31-OCT-1995; 95MO-EP004270.					
X	X					
P	R 01-NOV-1994; 94EP-00810627.					
X	X					
P	A (WELS/) WELS W.					
X	X					
P	I Weis W, Fominaya J;					
X	X					
D	R WP1, 1996-239505/24.					
X	X					
P	T P-PSDB; AAR35057.					
P	T Nucleic acid transfer system for gene therapy, e.g. against cancer - includes toxin translocation domain to target nucleic acid to specific cell.					
X	X					
S	Example 20, Page 84-86; 106pp; English.					

Pred. No.:	1,58e-94	Length:	19,9
Score:	1308.00	Matches:	245
Percent Similarity:	97.64%	Conservative:	3
Best Local Similarity:	96.46%	Mismatches:	6
Query Match:	55.78%	Indels:	0
DB:	2	Gaps:	0

US-09-596-774-7 (1-443) X AAT29412 (1-1919)

QY	1	GlnIleGlnIleuValGlnSerGlyProGluLeuValSerProGlyValThrValIle	20
Db	1147	CAGGTACACTGCACGACGTCTGGACCTTAAGTGAAGAGCTGGAGACAGTCAAGATC	1206
QY	21	SerCysIysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTrpValIysGlnAla	40
Db	1207	TCCTGCAGAGGCTCTGGGTATCTCTTTCACAAACTATGGAATGAACTGGGTGAAGAGCT	1266
QY	41	ProGlyGlnGlyLeuIleTyrPheMetGlyTrpIleAsnThrSerThrGlyIleSerThrPhe	60
Db	1267	CCAGAGACAGGGTTTAAAGTGAATGGAGTGGCTGAATTAAACCTCCACCTGAGAGATCAATTT	1326
QY	61	AlaAspAspPheIysGlyValArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr	80
Db	1327	GCTATGACTTCAAGAGGACGGTTGACTCTCTTGTGAAACCTTCGCCAACACTGCCTAT	1386
QY	81	LeuGlnIleAsnAsnLeuIleSerGlyAspPheAlaThrTyrPheCysAlaArgTrpGlu	100
Db	1387	TTGCAGATCAACAACCTCAAAAGTGAACAATGGCTCAATATTTCTGTGCANAGATGGAG	1446
QY	101	ValTyrHisGlyTyrValProTyrTrpGlyGlnGlyThrThrValThrValSerSerGly	120
Db	1447	GTTTACACACGGCTACGTTCTTACTCGGGGCCAAGGACACAGGTACCGGTTTCCCTGGC	1506
QY	121	GlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlyGlySerAspIleGlnLeuThrGln	140
Db	1507	GGTGGCGGTTCTGGTGGCGGTGGCTCCGGCGGTGGCTGTCTGCATCAGCTGACCCAG	1566
QY	141	SerHisIysPheLeuSerThrSerValGlyValAspArgValSerIleThrCysIysAlaSer	160
Db	1567	TTCTACAAATTCCTGTCACCTTCAGTGAAGAAGAGGAGTCAAGATCCTGCAGAGGCGCAGT	1626
QY	161	GlnAspValTyrAsnAlaValAlaTyrTrpGlnGlnIlySerProGlyGlnSerProIysLeu	180
Db	1627	CAGATGTGTATTAATGCTGTTGCCCTGGATACACAGAAACAGACATCTCTTAACCTT	1686
QY	181	LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly	200
Db	1687	CTGATTTACTGGGATCCTCCCGGAGACACTGGAATCCCTTCTCGGTTCACTGGCAGTGGC	1746
QY	201	SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr	220
Db	1747	TCTGGGCGCGGATTCACCTTACACATCAGCAGTGGCAGGCTGAAGACTTGAGCTTAT	1806
QY	221	PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyTyrThrLeuGluIle	240
Db	1807	TTCTGTACGACAACTTTTCGTACTTCATTCAGTTCCGCTCGGGGACAAATTTGAGAGATC	1866
QY	241	LysAlaLeuGluIleSerAsnSerValMetTyrPheSerSer 254	
Db	1867	AAAGCTCTAAGAGATCTCTCGAGTGAAGAGAAATTTTCAAGCC 1908	
RESULT 10			
AAT42038			
ID AAT42038 standard; DNA; 2214 BP.			
AC AAT42038;			
XX 29-JAN-1997 (first entry)			
XX Plasmid pMS238-5-TGF fragment encoding scFv (FRP5)/TGF alpha-ETA.			
XX Single chain antibody; scFv; monoclonal antibody; Mab; EGF; erbB-2;			
XX epidermal growth factor; receptor; plasmid pMS38-5-TGF; cancer;			
KW			

KW	exotoxin A; ETA; transforming growth factor alpha; TGF; antitumour; ss
XX	
OS	Homo; sapiens.
OS	Mus; sp.
OS	Pseudomonas; aeruginosa.
OS	Synthetic.
OS	Chimeric.
XX	
PH	Key
FT	CDS
FT	Location/Qualifiers
FT	1..2100
FT	/*tag= a
FT	1..63
FT	/*tag= b
FT	/note= "OmpA signal peptide"
FT	64..2097
FT	/*tag= c
FT	/product= "scFv(FRPs)/TGF alpha-ETA"

PN EP739984-A1.
XX
XX 30-OCT-1996.
PD
XX 26-APR-1995; 9SEP-00106275.
PF
XX 26-APR-1995; 9SEP-00106275.
PR 26-APR-1995; 9SEP-00106275.
XX
PA (SANT-) SAN TUMORFORSCHUNGS GMBH.
XX
XX
PI Wels W, Schmidt M, Groner B;
XX
DR WP1, 1996-478748/48.
P-PSDB; AAW05138.

PT Bivalent fusion proteins that bind epidermal growth factor receptor
 PT analogues - and comprise at least two different cell surface binding
 PT domain(s), useful for tumour therapy.
 XX
 Example 10; Page 25-27; 52pp; English.
 95

CC A from *Staphylococcus aureus* (GenBank accession number AF040308) and human
CC A from *Pseudomonas aeruginosa* and human transforming growth
CC factor (TGF- α) (see also AF040308). pMS238-5-TGF can be utilised in
CC the prodn. of a bispecific fusion protein in bacterial (esp. *E. coli*)
CC host cells. Such fusion proteins (see also AF040308-44) are useful as
CC antitumour agents

SQ Sequence 2214 BP; 431 A; 710 C; 665 G; 408 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1,88e-94	Length:	221
Score:	1308.00	Matches:	275
Percent Similarity:	71.198	Conservative:	24
Best Local Similarity:	65.488	Mismatches:	59
Query Match:	55.788	Indels:	64
DB:	2	Gaps:	5

US-09-596-774-7 (1-443) x AAT42038 (1-2214)

Qy	1	Gln116Gln156	Val141	InsertGly1	ProGlu156	Leu157	Ala158	Val159	Thr160	Val161	Val162	160
		115	CAAGTCAACTG	GACAGTCTG	GAACCTG	AACTG	GAAGAAAGCC	TGGAGAA	CAGTCAAG	ATC		176
Db		21	SeCysAla156	AsrGly157	ProPhe158	ThrAsn159	Trg156	Ala157	Thr158	Val159	Gln160	160
		175	TCTCGAAGAGCC	CTCGGATCT	CTTTCA	CAAACTA	TGAAATGA	AACTGGG	TGTAACAG	AGCT		233
Qy		41	ProGlu161	Gln162	Leu163	TrpMet164	Gly165	Thr166	Trn167	Glu168	Thr169	160
Db		235	CCAGGACAGGG	GTTTAAAGTGA	TGATGGG	CTGGATTA	CACTTCA	CTGAGAG	ATCAACAT	TTT		299


```

QY 61 AlaAspAspPheLeuYsgLYArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80
Db 295 GCTGATGACTTCAAGGACGGTTTACTTCTTGAAACCTCGCCACACCTGCTCAT 354
QY 81 LeuGluIleAsnAsnLeuLeuYsSerGluAspMetAlaThrTyrPheCysAlaArgTyrGlu 100
Db 355 TTGCAGATCAACAACTCAAGTAAAGACAGGCTCAATATTTCTGTGCAAGATGGAG 414
QY 101 ValTyrHisGlyTyrValProTyrTyrGlyGlnGlyThrThrValThrValSerSerGly 120
Db 415 GTTTCACACGGCTACTCTTCTTACTGGGGCCAAAGGACAGGTACCGTTCTCTG 474
QY 121 GlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThrGln 140
Db 475 GGTGGCGGCTTCTGGTGGCGGTGGCTCCGGCGGTGCTTGACATCCACTACCCAG 534
QY 141 SerHisGlyPheLeuSerThrSerValGlyAspArgValSerIleThrCysAlaAsp 160
Db 535 TCTCACAAAATTCCTGTCCACTTCAGTAGAGACAGGGTCACACATCACCTGCAGGCGCAGT 594
QY 161 GlnAspValTyrAsnAlaValAlaTyrTyrGlnGlnIlyProGlyGlnSerProGlyLeu 180
Db 595 CAGAGATGATTAATGCTGTTGCTGCTGATCAACAGAACAGACAAATCTCTAACTT 654
QY 181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200
Db 655 CTGATTTACTCGCATCTCCCGGTACACTGAGACCTCTTCTGCTTCACTGGCAGTGGC 714
QY 201 SerGlyProAspPheThrPheThrIleSerSerValGlnIleGluAspLeuAlaValTyr 220
Db 715 TCTGGGCGGATTCATCTTCACTCAACAGAGTGTGAGGCTGAAGCTGGGCGATTAT 774
QY 221 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyTyrThrLeuGluIle 240
Db 775 TTCTGTACAGCAACTTTTCTGATCCATCACTTCGGCTGGGGACAAATTTGCAATC 834
QY 241 LysAlaLeuGluIleSerAsnSerValMetTyrPheSerSerValProValLeuGln 260
Db 835 AAGGCTCTAGAGCA----- 848
QY 261 LysValAsnSerThrThrThrLysProValLeuArgThrProSerProValHisProThr 280
Db 849 -----CCATCATCACATCACTCAG 869
QY 281 GlyThrSerGlnProGlnArgProGluAspCysArgProArgIleSerValIysGlyThr 300
Db 870 GGG---CGGACGCTGGCGCGCGTAC-----CGCCACACAGGCTGCCACCTGCCGCT 920
QY 301 GlyLeuAspPheLeuGluAspProLysLeuCysTyrLeuLeuAspGlyIleLeuPheIle 320
Db 921 GGAGACTTTCAACCGTCATCGCGCGCGCGCTGGGAACAACTGGAGCAGTGGCGCTA 980
QY 321 TyrGlyValIle-IleThrAlaLeuTyrLeuArgAlaIlyPheSerArgSerAlaGluTh 340
Db 981 TCCGTTCCAGCGGTGGTGGCTCCCTTCACTGGCGCGCGCACTGTCA----- 1026
QY 340 rAlaAlaAsnLeuGlnAspProAsnGlnLeuTyrAsnGlnLeuAsnLeuGlyArgArgI 360
Db 1027 -----TGAACCAAGTGCACCA-GGT----- 1046
QY 360 uGluTyrAspValLeuGluIlyLysArgAlaArgAspProGluMetGlyIlyLysGlnG 380
Db 1047 -----GATCCCAACAGCGCTGGCCAGCCCGC 1072
QY 380 nArgArgArgAsnPro-GlnGluGlyValTyrAsnAlaLeuGlnIlyAspIlyMetAlaG 400
Db 1073 GCGAGCGCGCGACCTGGCGAGCGATCCGCGACAGCAGCCGAGCAGCGCGCTGCGCC 1132
QY 400 IuAlaTyrSerGluIleGlyThrLysGlyGluArgArgArgIlyLysGlnHisAsp 418
Db 1133 TGACCCGTGGCGCGCGCGAGAGCGAGCGCTTCCGTCCGCGAGGCGACCCGCGAACGAC 1188

```

RESULT 11

```

AAT42039
ID AAT42039 standard; DNA; 2793 BP.
XX AC AAT42039;
XX DT 29-JAN-1997 (first entry)
XX DE Plasmid pMS238-5-225 fragment encoding scFv2 (FRP5/225)-ETA.
XX KM Single chain antibody; scFv; monoclonal antibody; MAb; EGF; erbB-2;
XX KM epidermal growth factor; receptor; plasmid pMS238-5-225; cancer;
XX KM exotoxin A; ETA; antitumour; ss.
XX OS Mus; BP.
XX OS Pseudomonas; aeruginosa.
XX OS Synthetic.
XX OS Chimeric.
XX FH Key
XX FT CDS Location/Qualifiers
FT 1..2679
FT sig_peptide /+tag= a
FT 1..63
FT /note= "OmpA signal peptide"
FT 64..2676
FT /+tag= c
FT /product= "scFv2 (FRP5/225)-ETA"
XX FN EF39984-AL.
XX PD 30-OCT-1996.
XX PF 26-APR-1995; 95BP-00106275.
XX PR 26-APR-1995; 95BP-00106275.
XX (SANT-) SAN TUMORFORSCHUNGS GMBH.
XX PA Wels W, Schmidt M, Groner B;
XX PI WPI; 1996-478748/48.
XX DR P-PSDB; AAW05139.
XX PT Bivalent fusion proteins that bind epidermal growth factor receptor or
XX PT analogues - and comprise at least two different cell surface binding
XX PT domain(s), useful for tumour therapy.
XX PS Example 11; Page 28-30; 52pp; English.
XX CC A Fragment (AAT42039) of bacterial expression plasmid pMS238-5-225 codes
XX CC for scFv2 (FRP5/225)-ETA (AAW05139), and is obtd. by ligating DNA
XX CC fragments coding for the single-chain binding region of murine monoclonal
XX CC antibody FRP5 (see also AAT42034), portions of Pseudomonas aeruginosa
XX CC exotoxin A and the single-chain binding region of murine monoclonal
XX CC antibody 225 (see also AAT42033). pMS238-5-225 can be utilised in the
XX CC prodn. of bivalent fusion protein in bacterial (esp. E. coli) host cells.
XX CC Such fusion proteins (see also AAW05138-44) bind the epidermal growth
XX CC factor receptor and are useful as antitumour agents
XX SQ Sequence 2793 BP; 593 A; 847 C; 806 G; 547 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 2,496-94 Length: 2793
Score: 1108.00 Matches: 275
Percent Similarity: 71.1% Conservative: 24
Best Local Similarity: 65.4% Mismatches: 59
Query Match: 55.7% Indels: 64
DB: 2 Gaps: 5
US-09-596-774-7 (1-443) x AAT42039 (1-2793)
1 GlnIleGlnLeuValGlnSerGlyProGluLeuLysProGlyGluThrValIysIle 20
||||:|||||

```



```

Db 115 CAGGTACAACCTGACAGCTGACCTGAACTGAAAGCCTGAGAGACAGTCAAGATC 174
Qy 21 SerCysValSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValIleGlnIle 40
Db 175 TCCTGCAAGAGGCTCTGGGTATCTTTCACAACTATGAAATGAACTGGGTGAAGAGGCT 234
Qy 41 ProGlyGlnGlyLeuValTyrMetGlyTyrIleAsnThrSerThrGlyIleSerThrPhe 60
Db 235 CCAGAGACAGGCTTTAAATGATGGATGGCTGATTAACACTCCACCTGAGAGAGCAACTT 294
Qy 61 AlaAspAspPheLeuGlyArgPheAspPheSerLeuGlnThrSerAlaAsnThrAlaTyr 80
Db 295 GCTGATGACTTCAAGGAGCAGGTTTACTTCTTGGAAACCTGCAACACTGCTCAT 354
Qy 81 LeuGlnIleAsnAsnLeuLeuSerGlyLeuAspMetAlaTyrTyrPheCysAlaArgTyrGlu 100
Db 355 TTGCAGATCAACAACCTCAAAAGTAAAGACAGGCTCACTATTTCTGCAAGAGGAG 414
Qy 101 ValTyrHisGlyTyrValProTyrTyrPheGlyGlnGlyThrValThrValSerSerGly 120
Db 415 GTTATACACAGGCTACGTTCTTACTGAGGCGCAAGGAGCAGCGTTCCTCTGGC 474
Qy 121 GlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThrGln 140
Db 475 GGTGGCGGTTCTGGTGGCGGTGGCTCCGCGGTGGCTTGCATCAGCTGACCCAG 534
Qy 141 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysValAlaSer 160
Db 535 TCTCAACAATTCCTGTCACCTCAGTAGAGACAGGAGCAGATCACTGCAAGGCCAGT 594
Qy 161 GlnAspValTyrAsnAlaValAlaTyrPheGlnGlnLysProGlyGlnSerProLysLeu 180
Db 595 CAGAGTGTATATATGCTGTTGCTGATCAACAGAAACAGACATCTCTTAATCT 654
Qy 181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200
Db 655 CTGATTTACTCGGCATCTCCCGGTACACTGAGAGTCCCTCTCGCTTCACTGCGAGTGGC 714
Qy 201 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 220
Db 715 TCTGGGCGGATTTCACTTCAACATCAGAGTGGAGGCTGAGACCTGCGAGTTTAT 774
Qy 221 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyThrLysLeuGlnIle 240
Db 775 TTCGTGACGACAATTTTCGATCTCATTCACTTCGCTCGGGGACAAAATTGGAGATC 834
Qy 241 LysAlaLeuGlnIleSerAsnSerValMetTyrPheSerSerValProValLeuGln 260
Db 835 AAAGCTCTAGAGCA----- 848
Qy 261 LysValAsnSerThrThrThrLysProValLeuArgThrProSerProValHisProThr 280
Db 849 -----CCATCATCACTCACTCACTGAGA 869
Qy 281 GlyThrSerGlnProGlnArgProGlnAspCysArgProArgGlySerValLysGlyThr 300
Db 870 GGG---CGGCAAGCTGGCGCGCTGAC-----CGGCAACAGGCTGCTCCACCTGCGCT 920
Qy 301 GlyLeuAspPheLeuGlnAspProLysLeuGlyCysTyrIleLeuAspGlyIleLeuPheIle 320
Db 921 GGAGACTTTTCAACCGTCATCGCCAGCGCGCGCTGGGAGCAACTGAGAGCAGTGGGGCTA 980
Qy 321 TyrGlyValIle-IleThrAlaLeuTyrLeuArgAlaLysPheSerArgSerAlaGluThr 340
Db 981 TCCGCTGACAGCGGCTGCTGCTCTTACTGACGCGCGGAGCTGCA----- 1026
Qy 340 rAlaAlaAsnLeuGlnAspProAsnGlnLeuTyrAsnGlnLeuAsnLeuGlnIleArgGln 360
Db 1027 -----TGGAACCAAGCTCAACA-GGT----- 1046
Qy 360 uGluTyrAspValLeuGlnLysLysArgAlaArgAspProGlnLysGlyGlnIle 380
Db 1047 -----GATCCGCAAGCGCCTGGCCAGCCCG 1072

```

```

Qy 380 nArgArgAsnPro-GlnGluGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaG 400
Db 1073 GCAGCGCGCGGAGCTGGGCGGAGGATCCCGAGACCGCGAGAGGCCCTGTGGGCC 1132
Qy 400 IuAlaTyrSerGluIleGlyThrLysGlyGlnArgArgGlyLysGlyHisAsp 418
Db 1133 TGACCTGCGCGCGCGGAGGAGCGAGCGCTTGTCCGCGAGGCGCACCGCAAGCAG 1188

RESULT 12
AAT42043
ID AAT42043 standard; DNA: 2793 BP.
AC AAT42043;
DT 29-JAN-1997 (first entry)
DE Plasmid pMS238-5-5 encoding scFv2 (FRP5/FRP5)-ETA (version 2).
XX
KW Single chain antibody; scFv; monoclonal antibody; MAb; EGF; erbB-2;
KW epidermal growth factor; receptor; plasmid pMS238-5-5; cancer;
KW exotoxin A; ETA; antitumour; ss.
XX
OS Mus; sp.
OS Pseudomonas; aeruginosa.
OS Synthetic.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT CDS 1..2679
FT /*tag= a
FT sig_peptide 1..63
FT /*tag= b
FT /note= "Ompa signal peptide"
FT mat_peptide 64..2676
FT /tag= c
FT /product= "scFv2 (FRP5/FRP5)-ETA"

XX
BN BP739984-A1.
XX
PD 30-OCT-1996.
XX
PF 26-APR-1995; 95EP-00106275.
XX
PR 26-APR-1995; 95EP-00106275.
XX
PA (SANT-) SAN TUMORFORSCHUNGS GMBH.
XX
PI Weis W, Schmidt M, Groner B;
XX
DR WP1; 1996-478748/48.
DR P-PSDB; AAM05143.
XX
PT Bivalent fusion proteins that bind epidermal growth factor receptor or
PT analogues - and comprise at least two different cell surface binding
PT domain(s), useful for tumour therapy.
XX
PS Example 12; Page 40-42; 52PP; English.
XX
CC A fragment (AAT42043) of bacterial expression plasmid pMS238-5-5 codes
CC for scFv2 (FRP5/225)-ETA (AAM05143), and is obtd. by ligating 2 copies of
CC DNA coding for the single-chain binding region of murine monoclonal
CC antibody FRP5 (see also AAT42034) and portions of Pseudomonas aeruginosa
CC exotoxin A. pMS238-5-5 can be utilised in the prodn. of bivalent fusion
CC protein in bacterial (esp. E. coli) host cells. Such fusion proteins (see
CC also AAM05138-44) bind the epidermal growth factor receptor and are
CC useful as antitumour agents
XX
SO Sequence 2793 BP; 578 A; 858 C; 812 G; 545 T; 0 U; 0 Other;

Alignment Scores: 2,49e-94 Length: 2793
Pred. No.: 1308.00 Matches: 275
Score:

```

Percent Similarity: 71.19%
 Best Local Similarity: 65.48%
 Query Match: 55.78%
 DB: 2
 Conservative: 24
 Mismatches: 59
 Indels: 64
 Gaps: 5

US-09-596-774-7 (1-443) x AAT42043 (1-2793)

```

OY 1 GlnIleGlnLeuValGlnSerGlyProGluLeuLysLysProGluValThrValIle 20
DB 115 CAGGTACAACTGCAGCACTGTCGACTGAACTGAAAGAGCCCTGGAGAGACGTCAAGATC 174
OY 21 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValIleGlnAla 40
DB 175 TCCTGCAGAGCCCTCGGGTATCCTTTCACAACTAGTAAGAACTGGGTGAAGAGAGCT 234
OY 41 ProGlyGlnGlyLeuLysTyrMetGlyTyrIleAsnThrSerThrGlyLysSerThrPhe 60
DB 235 CCAGAGACAGGGGTTTAAAGTGAAGGCGGTGAATTAACACTCCACTGAGAGTCAACATTT 294
OY 61 AlaAspAspPheLysGlyArgPheAspPheSerLeuGlnThrSerAlaAsnThrAlaTyr 80
DB 295 GCTGATGACTTCAGAGGAGCGGTTTGACTTCTTGGAACTCTGCGCAACACTGCTAT 354
OY 81 LeuGlnIleAsnAsnLeuLysSerGluAspMetAlaThrTyrPheCysAlaArgTyrGln 100
DB 355 TTGCAGATCAACAACCTCAAAAGTGAAGACATGGCTACATATTTCTGTGCAAGATGGAG 414
OY 101 ValTyrHisGlyTyrValProTyrTyrGlyGlnGlyThrValThrValSerSerGly 120
DB 415 GTTACCAACGGCTACGCTTCTTACTGGGGGCAAGAGGACACGGTACCGCTTCTGTC 474
OY 121 GlyLysGlySerGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThrGln 140
DB 475 GGTGGCGGTTTGTGGCGGTGGCTCCGCGCGGTGGCTTGTGACATCCAGTCAACCCAG 534
OY 141 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSer 160
DB 535 TCTCACAATTCCTGCTCCTCACTTCAGTAGAGACAGGCTCAGCATCACCTGCAAGCCAGT 594
OY 161 GlnAspValTyrAsnAlaValAlaTyrGlnGlnLysProGlyGlnSerProLysLeu 180
DB 595 CAGAGATGTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 654
OY 181 LeuIleTyrSerAlaSerSerSerTyrTyrGlyValProSerArgPheThrGlySerGly 200
DB 655 CTGATTAATCTCGCATCTCCCGGTACACTGAGTCTCTTCTGCTTCACTGCGCAGTGGC 714
OY 201 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 220
DB 715 TCTGGGCGGATTTTCACTTTCACATCAGCAGTGTGAGCGCTGAAGACTGGCAGTTTAT 774
OY 221 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyThrLysLeuGlnIle 240
DB 775 TTCGTGTACGACAACTTTTGTGACTTCATTCAGTTCGGCTCGGGGACAAATTTGGAGATC 834
OY 241 LysAlaLeuGlnIleSerAsnSerValMetTyrPheSerSerValProValLeuGln 260
DB 835 AAAGCTCTAGAGCA----- 848
OY 261 LysValAsnSerThrThrLysProValLeuArgThrProSerProValHisProThr 280
DB 849 -----CCATCATACCATCATCAGTCA 869
OY 281 GlyThrSerGlnProGlnArgProGluAspCysArgProArgLysValLysGlyThr 300
DB 870 GGG---CGGAGCTTGGCGCGCGTAC-----CGCGACACAGGCTGCCACCTGCCGCT 920
OY 301 GlyLeuAspPheLeuGlnAspProLysLeuCysTyrLeuLeuAspGlyIleLeuPheIle 320
DB 921 GGAGACTTTTACCGCGTATCGCGCAGCGCGGCTGGGAACAACCTGAGACAGTGGCGCTA 980
OY 321 TyrGlyValIle-IleThrAlaLeuTyrLeuAlaGlyAlaLysPheSerArgSerAlaGlu 340
DB 340 ----- 340

```

```

DB 981 TCCGGTGACAGCGGTGTCGCGCTTCACTGCGCGCCGCACTGTCA----- 1026
OY 340 rAlaAlaAsnLeuGlnAspProAsnGlnIleuTyrAsnGlnLeuAsnLeuGlyArgArgI 360
DB 1027 -----TGGAACAGAGTGACCA--GGT----- 1046
OY 360 uGluTyrAspValLeuGlnLysLysArgAlaArgAspProGluMetGlyLysGlnI 380
DB 1047 -----GATCCGCAACGCCCTGGCCAGCCCCCG 1072
OY 380 nArgArgArgAsnPro-GlnGlnGlyValTyrAsnAlaLeuGlnLysAspLysMetAla 400
DB 1073 CGAGCGCGCGGCACTCGCGGAGCGATCCGCGAGACGCCGAGACAGGCCCGTGTGCC 1132
OY 400 LuAlaTyrSerGlnIleGlyThrLysGlyGluArgArgArgGlyLysGlyHisAsp 418
DB 1133 TGACCTTGCGCGCGCGCGAGGAGGCGCTTGTGTCGCGAGCGGACCGCAACGAC 1188

RESULT 13
AAT42041
ID AAT42041 standard; DNA; 3177 BP.
XX
AC AAT42041;
XX
DT 29-JAN-1997 (first entry)
XX
DE Plasmid pMS240-5-225 fragment encoding scFv2 (FRP5/225)-ETA.
XX
KW Single chain antibody; scFv; monoclonal antibody; MAbs; EGF; erbB-2;
KW epidermal growth factor; receptor; plasmid pMS240-5-225; cancer;
KW exotoxin A; ETA; antitumour; ss.
XX
OS Mus; sp.
OS Pseudomonas; aeruginosa.
OS Synthetic.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT CDS 1..3063
FT sig_peptide /tag= a
FT /tag= b
FT /note= "OmpA signal peptide"
FT mat_peptide 64..3061
FT /tag= c
FT /product= "scFv2 (FRP5/225)-ETA"

XX EP739984-A1.
XX
XX 30-OCT-1996.
XX
XX 26-APR-1995; 95EP-00106275.
XX
XX 26-APR-1995; 95EP-00106275.
XX
XX 26-APR-1995; 95EP-00106275.
XX
XX (SANT-) SAN TUMORFORSCHUNGS GMBH.
XX
XX Wels W, Schmidt M, Groner B;
XX
XX WPI; 1996-4/78/48/48.
XX
XX P-PSDB; AAM05141.
XX
XX Bivalent fusion proteins that bind epidermal growth factor receptor or
XX PT analogues - and comprise at least two different cell surface binding
XX PT domain(s), useful for tumour therapy.
XX
XX Example 11; Page 34-36; 52pp; English.
XX
XX A fragment (AAT42041) of bacterial expression plasmid pMS240-5-225 codes
XX for scFv2 (FRP5/225)-ETA (AAM05141), and is obtd. by ligating DNA
XX fragments coding for the single-chain binding region of murine monoclonal
XX antibody FRP5 (see also AAT42034), portions of Pseudomonas aeruginosa
XX exotoxin A and the single-chain binding region of murine monoclonal

```

CC antibody 225 (see also AAT42033). pMS240-5-225 can be utilized in the
 CC prodn. of bivalent fusion protein in bacterial (esp. E. coli) host cells.
 CC Such fusion proteins (see also AAW05138-44) bind the epidermal growth
 CC factor receptor and are useful as antitumour agents

XX Sequence 3177 BP, 654 A, 993 C, 939 G, 591 T, 0 U, 0 Other;

Alignment Scores:

Pred. No.:	2,91e-94	Length:	3177
Score:	1308.00	Matches:	275
Percent Similarity:	71.19%	Conservative:	24
Best Local Similarity:	65.48%	Mismatches:	59
Query Match:	55.78%	Indels:	64
DB:	2	Gaps:	5

US-09-596-774-7 (1-443) x AAT42041 (1-3177)

QY	1	GlnIleGlnLeuValGlnSerGlyProGluLeuLysLysProGlyGluThrValIle	20
DB	115	CAGGTACAACTGCAGCAGTCTGACCTGAACTGAAGAACTGGAGAGACAGTCAAGATC	174
QY	21	SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTrpValIleGlnAla	40
DB	175	TCCGTCAAGGCGCTCTGGGTATCTTTCCAAACTATGAAATGAACCTGGGTGACAGGCT	234
QY	41	ProGlyGlnGlyLeuLysTrpMetGlyTyrIleAsnThrSerThrGlyLysSerThrPhe	60
DB	235	CCAGGACAGGGGTTTAAAGTGGATGGCTGAGATTAACCTTCACCTGGAGATCAACATT	294
QY	61	AlaAspAspPheLysGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr	80
DB	295	GCTGATGACTTCAAGGAGCGGTTGACTTCTCTTGAAACCTGCGCAACAGCGCTAT	354
QY	81	LeuGlnIleAsnAsnLeuLysSerGluAspMetAlaThrTyrPheCysAlaArgTrpGlu	100
DB	355	TTGGAGATCAACAACTCAAAAGTGAAGACATGGCTCATTTCTGTGCAAGATGGAG	414
QY	101	ValTyrHisGlyTyrValProTyrTrpGlyGlnGlyThrValThrValSerSerGly	120
DB	415	GTTTACCAAGGCTTACGTTCTTACTGGGCGCAAGGACACGCTGCTCTGGC	474
QY	121	GlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThrGln	140
DB	475	GGTGGCGGTTCTGGTGGCGGTCCTCGCGCGGTTCTGACATCCAGTACCAG	534
QY	141	SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSer	160
DB	535	TCTCACAAAATTCCTGCTCAGTTCAGTAGAGACAGGCTCAGCATCAGCTGCAGGCA	594
QY	161	GlnAspArgTyrAsnAlaValAlaTrpTyrGlnGlnLysProGlyGlnSerProLysLeu	180
DB	595	CAGATGTGTATTAATGCTGTGCTGTGATCAACAGAAACAGGACAAATCTCTAACTT	654
QY	181	LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly	200
DB	655	CTGATTTACTCGGATCTCTCCCGGTACACGTGATCTCTCTCTCAGTGGAGTGGC	714
QY	201	SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr	220
DB	715	TCTGGGCGGATTTCACTTTCACATCAGCAGTGCAGGCTGAAAGACTGGCAGTTAT	774
QY	221	PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyTyrThrLysLeuGluIle	240
DB	775	TTCTGTAGCAAACTTTTTCGATCTCAATCAGCTTCGCTGGGACAAATTTGAAATC	834
QY	241	LysAlaLeuGluIleSerAsnSerValMetTyrPheSerSerValProValLeuGln	260
DB	835	AAAGCTTACAGCA-----	848
QY	261	LysValAsnSerThrThrLysProValLeuArgThrProSerProValHisProThr	280
DB	849	-----CCATCATCAGCATCAGCTCAGA	869

QY	281	GlyThrSerGlnProGlnArgProGluAspCysArgProArgIleSerValIleGlyThr	300
DB	870	GGG---CGGACGCTTGCGCGCTGAC-----CGGACACAGCGCTGCCACCTGGCGCT	920
QY	301	GlyLeuAspPheLeuGluAspProLysLeuCysTyrLeuLeuAspGlyIleLeuPheIle	320
DB	921	GGACACTTTCACCCGTCATCGCCCGCGCGCTGGGAACACTGAGCAGTCCGGCTA	980
QY	321	TyrGlyValIle-IleThrAlaLeuTyrLeuArgAlaLysPheSerArgSerAlaIle	340
DB	981	TCCGTGACAGCGGCTGTGCTGCTTCTACCTGGCGCGGACCTGTA-----	1026
QY	340	AlaAlaAsnLeuGlnAspProAsnGlnLeuTyrAsnGlyLeuAsnLeuGlyArgArgGly	360
DB	1027	-----TGAAACAGGTGACCA-GGT-----	1046
QY	360	uGluTyrAspValLeuGluLysLysArgAlaArgAspProGluMetGlyIleGlnGlnI	380
DB	1047	-----GATCCGCAACGCCCTGGCCAGCCCG	1072
QY	380	ArgArgArgAspPro-GlnGluGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaG	400
DB	1073	GCAAGCGCGGCGACTGCGGCAACGATCGCGACAGACCGGAGACGCGCTGCGCC	1132
QY	400	IuAlaTyrSerGluIleGlyThrLysGlyGluArgArgGlyLysGlyHisAsp	418
DB	1133	TGACCTTGGCGCGCGGACGAGCGGCTTCTGCGGACGAGCAGCGGCAACAG	1188

RESULT 14
 AAT29409
 ID AAT29409 standard; DNA, 1692 BP.
 XX
 AC AAT29409;
 XX
 DT 16-OCT-2003 (revised)
 DT 18-AUG-1996 (first entry)
 XX
 DE scFv(FRPS)-DETA-DGAL4 fusion gene in pW46-5.
 XX
 KW Nucleic acid transfer system; gene transfer; gene therapy;
 KW cell targeting; multidomain protein; vector; cancer; exotoxin A; DETA;
 KW ompA; signal peptide; single chain antibody; scFv; Gal4; ss.
 XX
 OS Escherichia; coli.
 OS Mus; sp.
 OS Pseudomonas; aeruginosa.
 OS Saccharomyces cerevisiae.
 OS Chimeric.
 OS Synthetic.
 XX
 FH Key
 FT CDS
 FT Location/Qualifiers
 FT 1..1656
 FT /*tag= a
 FT s1g_peptide
 FT /*tag= b
 FT mat_peptide
 FT 64..1653
 FT /*tag= c
 FT /product= "scFv(FRPS)-DETA-DGAL4"
 FT
 XX
 PN W09613599-A1.
 XX
 PD 09-MAY-1996.
 XX
 PF 31-OCT-1995; 95W0-EP004270.
 XX
 PR 01-NOV-1994; 94EP-00810627.
 XX
 PA (WELLS/) WELLS W.
 XX
 FI Wells W, Fominaya J;
 XX
 DR WPI; 1996-239505/24.
 DR P-PSDB; AAR35053.


```

FT      /*tag= d
FT      /label= Linker
FT      misc_RNA      491..814
FT      /*tag= e
FT      /label= FRP5_light_chain_variable_domain
FT      mat_peptide      815..2155
FT      /*tag= f
FT      /label= phoA_coding_region
XX      EP502812-A1.
XX      09-SEP-1992.
XX      27-JAN-1992; 92EP-00610056.
XX      05-FEB-1991; 91EP-00810079.
XX      (CIBA ) CIBA GEIGY AG.
XX      Wels WS, Hynes NE, Harwerth I, Groner B, Hardman N, Zwickl M;
XX      WPI; 1992-302096/37.
XX      P-PSDB; AAR26980.
XX      Recombinant antibodies directed to growth factor receptor C-ERBB-2 - for
XX      diagnosing and treating tumours expressing C-ERBB-2 e.g. breast or
XX      ovarian tumours.
XX      PS      Disclosure; Page 34-40; 67pp; English.
XX      CC      The sequence given encodes the single chain recombinant antibody
XX      CC      designated Fv(FRP5)-phoA. The alkaline phosphatase gene (phoA) was used
XX      CC      as a marker gene so that E. coli transformed with the fusion gene could
XX      CC      be identified. The fusion gene was expressed in E. coli and the antibody
XX      CC      was extracted. This recombinant antibody can be used for the qualitative
XX      CC      and quantitative determination of c-erbB-2. This can be used for
XX      CC      monitoring or in-vivo localisation of tumours overexpressing c-erbB-2.
XX      CC      (Updated on 25-MAR-2003 to correct PN field.)
SQ      Sequence 2233 BP; 566 A; 573 C; 606 G; 488 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:      4,74e-94      Length:      2233
Score:      1303.00      Matches:      245
Percent Similarity:      95.00%      Conservative:      2
Best Local Similarity:      94.23%      Mismatches:      1
Query Match:      55.57%      Indels:      12
DB:      2      Gaps:      1

US-09-596-774-7 (1-443) x AAQ28257 (1-2233)
QY      1  GlnlleglneulvalGlnSerGlyProGluLeuLysProGlyGluThrValLysIle 20
DB      89  CAGGTACCACTGCAGCAGTCTGCACCTGAACCTGAAGAGCTGGAGAGACAGTCAAGATC 148
QY      21  SerGlyValAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTrpValLysGlnAla 40
DB      149  TCCGCAAGGCGCTCTGGGTATCTTTCACAACTATGGAATGAACCTGSGTGAAGCAGGCT 208
QY      41  ProGlyGlnGlyLeuLysTyrPheMetGlyTyrPheLeuAsnThrSerThrGlyLysSerThrPhe 60
DB      209  CCAGAGACAGGCGTTTAACTGATGGCGTGAATTAACACCTCCACCTGAGAGATCAACATTT 268
QY      61  AlaAspAspPheLysGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80
DB      269  GCTGATGACTTCAAGGAGCGGTTTGACTTCTTTGGAAACCTCGCCAAACACTGCGCTTAT 328
QY      81  LeuGlnIleAsnAsnLeuLysSerGluAspMetAlaThrTyrPheCysAlaArgTrpGlu 100
DB      329  TTGCAGATCAACAACCTCAAAAGTGAAGACATGGCTACATATTTCTGTGCAAGATGGAG 388
QY      101  ValTyrHisGlyTyrValProTyrTrpGlyGlnGlyThrThrValThrValSerSerGly 120

```

```

DB      389  GTTTACACAGGCTTACTCTTCTGAGGCGCAAGGACCAACGATCACCCTTCTCTGCG 448
QY      121  GlyGlySerGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThrGln 140
DB      449  GGTGGCGGTTCTGTGGCGGTGGCTCCGGCGGTGGCTTGTGACATCAGCTAACCCAG 508
QY      141  SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysAlaAsp 160
DB      509  TCTCACAATTCCTGTCACATTCAGTAGGAGACAGGGGTCAACATCCTGCAAGGCCAGT 568
QY      161  GlnAspValTyrAsnAlaValAlaTyrGlnGlnLysProGlyGlnSerProLysLeu 180
DB      569  CAGGATGTATTAATGCTGTGGCTGGTATCAACAGAAACCAAGACAAATCTTAAACTT 628
QY      181  LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200
DB      629  CTGATTTACTCGGCATCTCCGGGTACAGTGAAGTCCCTTCCTCCTCAGTGGCAGTGGC 688
QY      201  SerGlyProAspPheThrPheThrIleSerSerValGlnIleAspLeuAlaValTyr 220
DB      689  TCTGGCGCGATTTCACTTTCACATCAGCAGGTGTGACAGGCTGAAGACTGGCAGTTTAT 748
QY      221  PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyThrLysLeuGluIle 240
DB      749  TTCTGTGAGCAACATTTTCGTACTTCATTCACGTTCCGCTCGGGGACAAAATTGGAGATC 808
QY      241  LysAlaLeuGlnIleSerAsnSerValMetTyrPheSerSerValProValLeuGln 260
DB      809  AAAGCTCTAGAG-----CCTGTTCTGGAA 832

```

Search completed: April 20, 2005, 10:44:55
 Job time : 737.644 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005. Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 20, 2005, 10:19:14 ; Search time 224.44 Seconds
(without alignments)
3229.682 Million cell updates/sec

Title: US-09-596-774-7

Perfect score: 2345
Sequence: 1 QIQVQSGPELKKPGETVKI.....LSTAKDTYDALHQTIAPR 443

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 81813835 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=x1h
-Q=/cgn2_1/USPRO.spool_h/US09596774/runat_20042005_075443_12533/app_query.fasta_1.1220
-DB=Issued_Patents_NA -QPMT=fastcap -SUFFIX=aprx20.rml -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MUTLEN=0 -MAXLEN=2000000000
-USPR=US09596774.QCGN_1.1_93 @runat_20042005_075443_12533 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERRY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6
-Fgapext=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTUS.COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Prod. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1310.5	55.9	2012	1	US-08-235-838-13
2	1310.5	55.9	2012	2	US-08-465-473B-13
3	1308	55.8	1862	4	US-08-840-713-36
4	1308	55.8	1919	4	US-08-840-713-34
5	1305	55.7	1692	4	US-08-840-713-1
6	1303	55.6	2233	1	US-08-235-838-6
7	1303	55.6	2233	2	US-08-465-473B-6
8	1285.5	54.8	748	1	US-08-235-838-4
9	1285.5	54.8	748	2	US-08-465-473B-4
10	984.5	42.0	1869	2	US-08-356-786-15
11	981.5	41.9	909	1	US-08-133-804-1
12	981.5	41.9	909	1	US-08-461-184-7

13	981.5	41.9	909	1	US-08-463-675-7	Sequence 7, Appli
14	981.5	41.9	909	1	US-08-464-589-7	Sequence 7, Appli
15	981.5	41.9	909	1	US-08-461-838-1	Sequence 1, Appli
16	981.5	41.9	909	2	US-08-461-386-1	Sequence 1, Appli
17	903.5	38.5	723	1	US-07-988-430-93	Sequence 93, Appli
18	903.5	38.5	723	1	US-08-425-336-90	Sequence 90, Appli
19	903.5	38.5	723	1	US-08-468-113B-90	Sequence 90, Appli
20	903.5	38.5	723	1	US-08-477-484B-90	Sequence 90, Appli
21	903.5	38.5	723	2	US-08-646-360-90	Sequence 90, Appli
22	903.5	38.5	723	3	US-08-839-765-80	Sequence 90, Appli
23	903.5	38.5	723	3	US-09-136-389-90	Sequence 90, Appli
24	903.5	38.5	723	3	US-09-610-838-90	Sequence 90, Appli
25	903.5	38.5	723	4	US-09-711-485-90	Sequence 90, Appli
26	903.5	38.5	723	5	PCT-US92-09487-93	Sequence 93, Appli
27	892.5	38.1	1098	3	US-08-875-811-54	Sequence 54, Appli
28	890	38.0	1095	3	US-08-875-811-52	Sequence 52, Appli
29	889.5	37.9	1626	3	US-09-485-737B-84	Sequence 84, Appli
30	889.5	37.9	1626	4	US-10-071-485-84	Sequence 84, Appli
31	888.5	37.9	804	3	US-09-485-737B-1	Sequence 1, Appli
32	888.5	37.9	804	4	US-10-071-485-1	Sequence 1, Appli
33	888.5	37.9	2133	3	US-09-485-737B-89	Sequence 89, Appli
34	888.5	37.9	2133	4	US-10-071-485-89	Sequence 89, Appli
35	884.5	37.7	883	3	US-09-184-658-7	Sequence 7, Appli
36	884.5	37.7	883	4	US-09-504-262D-7	Sequence 7, Appli
37	878.5	37.5	1605	2	US-08-356-786-9	Sequence 9, Appli
38	866.5	37.0	786	3	US-08-635-928-11	Sequence 31, Appli
39	854	36.4	777	4	US-10-092-246-5	Sequence 5, Appli
40	854	36.4	777	4	US-10-092-246-7	Sequence 7, Appli
41	854	36.4	777	4	US-10-096-246A-5	Sequence 5, Appli
42	854	36.4	777	4	US-10-096-246A-7	Sequence 7, Appli
43	848	36.2	702	3	US-09-485-737B-92	Sequence 92, Appli
44	848	36.2	702	4	US-10-071-485-92	Sequence 92, Appli
45	846	36.1	777	4	US-10-092-246-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-08-235-838-13
Sequence 13, Application US/08235838
Patent No. 5571894
GENERAL INFORMATION:
APPLICANT: Wells, Manfred S.
APPLICANT: Hynes, Nancy E.
APPLICANT: Harwerth, Ina-Maria
APPLICANT: Groemer, Bernd
APPLICANT: Hartman, No. 5571894man
TITLE OF INVENTION: Zwickl, Markus
TITLE OF INVENTION: Recombinant Antibodies Specific for a
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: New York
COUNTRY: USA
ZIP: 10532
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235,838
FILING DATE: TBA
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/828,832
FILING DATE: 31-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 91-810079.3
FILING DATE: 05-FEB-1991

ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: 4-18518/A/CLP/CONT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919)541-8614
TELEFAX: (919)541-8689
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 2012 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Mouse/Pseudomonas aeruginosa
INDIVIDUAL ISOLATE: E. coli
IMMEDIATE SOURCE:
CLONE: pMW215-5
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1911 /note= "64-87 FLAG peptide and
OTHER INFORMATION: enterokinase cleavage site; 97-453 heavy chain
OTHER INFORMATION: variable domain;454-498 15 aa linker sequence;
OTHER INFORMATION: 499-822 FRPS light chain variable domain
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 1..63
OTHER INFORMATION: /note= "ompA signal peptide"
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 94..1911
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 1912..2012
OTHER INFORMATION: /function= "3'non-coding region of
OTHER INFORMATION: the exotoxin A gene"
FEATURE:
NAME/KEY: misc_feature
LOCATION: 826..1911
OTHER INFORMATION: /note= "Exotoxin A gene coding
OTHER INFORMATION: region (coding for amino acids 252 to 613 of the
OTHER INFORMATION: mature exotoxin A)"
US-08-235-838-13
Alignment Scores:
Pred. No.: 9,37e-117 Length: 2012
Score: 1310.50 Matches: 279
Percent Similarity: 68.51% Conservative: 19
Best Local Similarity: 64.14% Mismatches: 43
Query Match: 55.88% Indels: 95
DB: 1 Gaps: 16
US-09-596-774-7 (1-443) x US-08-235-838-13 (1-2012)
QY 1 GlnIleGlnLeuValGlnSerGlyProGluLeuLysProGlyGluThrValLysIle 20
DB 97 CAGGTACAACTGACGAGCTGACCTGAACTGAAAGCCTGGAGAGACAGTCAAGATC 156
QY 21 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnThrValLysGlnAla 40
DB 157 TCCTGCAAGCCCTCTGGGTATCCTTTCAAACTATGAACTGAATGAACTGGGTGAAGCAAGCT 216
QY 41 ProGlyGlnGlyLeuLysTyrPheMetGlyTyrIleAsnThrSerThrGlyGluSerThrPhe 60
DB 217 CCAGGACAGAGGTTAAAGTGAATGAGCTGATTAACACCTCCACTGAGAGTCAACATTT 276
QY 61 AlaAspAspPheLysGlyTyrArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80
DB 277 GCTGATGACTTCAAGGAGCGTTGACTTCTTTGGAAACCTCTGCAACACTGCTAT 336

QY 81 LeuGlnIleAsnLeuLysSerGlyAspMetAlaThrTyrPheCysAlaAgtProGlu 100
DB 337 TTGAGATCAACAACCTCAAAAGTGAACATGGCTCAATTTCTGTGCAAGATGGAG 336
QY 101 ValTyrHisGlyTyrValProTyrTyrGlyGlnGlyThrThrValThrValSerSerGly 120
DB 397 GTTACCAAGCGCTACGTTCTTACTGGGAGCCAGAGACCGATCCGTTCTCTGCGC 456
QY 121 GlyGlyGlySerGlyGlyGlySerGlyGlyGlyGlySerAspIleGlnLeuThrGln 140
DB 457 GGTGGCGGTTCTGGTGGCGGTGGCTCCGGCGGTGGCGGTCTGACATCCAGCTGACCCAG 516
QY 141 SerHisLysPheLeuSerThrThrSerValGlyAspArgValSerIleThrCysLysAlaSer 160
DB 517 TCTCAAAATTCCTGTCACCTCAGTAGAGACAGAGGTCCAGTACCTCCAGAGCCAGT 576
QY 161 GlnAspValTyrAsnAlaValAlaTyrTyrGlnGlnLysProGlyGlnSerProLysLeu 180
DB 577 CAGGATGTGTAATAGTGTGCTGCTGGTATCAACAGAAACAGACATCTCTTAACCT 636
QY 181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200
DB 637 CTGATTTACTCGGCATCTCCCGGTACCTGAGATCCCTTCTGCTTCACTGGCAGTGGC 696
QY 201 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 220
DB 697 TCTGGCGCGGATTTCACTTTCACCATCAGCGTCCAGCGTGAACCTGGCAGTTTAT 756
QY 221 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyThrLysLeuGluIle 240
DB 757 TTCTGTACAGCAATTTTGTCTACTCATTCACCTTGGCTCGGGAGCAAAATTTGGAGATC 816
QY 241 LysAlaLeuGlu-----IleSerAsnSerValMetTyrPheSerSerValPro 257
DB 817 AAAGCTCTAGAGGCGCGCAGCCTGCGCGCGCTGACCCGACCCAGCGCTGCCACCTGCCG 876
QY 258 ValLeuGlnLysValAsnSerThrThrThrLys----- 268
DB 877 CTG-----GAGACTTTCACCCGCTCATCGCCAGCCGCGCGCTGGGAACA 921
QY 269 -----ProValLeuArgThrProSerProValHisProThrGlyThr 282
DB 922 CTGAGACAGTGGCGCTATCCGGTGCAGCGCGTGGT--CGCCCTCTA---CTGGCGCGCGG 977
QY 283 Ser-----GlnProGlnArgProGluAspCysArgProArgLysSerValLysGlyThr 300
DB 978 ACTGTCATGGAACACAGTCGACCGACGAGTGAATCCGCAACGCCCTGGCCAGCCCGGACGCG 1037
QY 301 GlyLeuAspPheLeuGlnLysProLysLeuCySTyrLeuLeuAspGlyIleLeuPheIle 320
DB 1038 CGGCGACCTGGGCGGAACCAATCCG----- 1061
QY 321 TyrGlyValIleIleThrAlaLeuTyrIleuArgAlaLysPheSerArgSerAlaGluThr 340
DB 1062 -----CGAGCAGCCGAGAGA 1076
QY 341 AlaAlaAsnLeuGlnAspProAsnGlnLeuTyrAsnGlnLeuAsnLeuGlyArgGlu 360
DB 1077 GGCCCTCTGGCCCTGACCT-----GGCCCGCGCGCA 1109
QY 361 GluTyrAspValLeuGlnLysLysArgAlaArgAspProGluMetGlyGlyLysGlnGln 380
DB 1110 GAG-----CGAGCGCTTCTGCTCG-----CGAGGGGCAACGGGCA 1142
QY 381 ArgArgArgAsnProGlnGlnLysValTyrAsnAlaLeuGlnLysAspLysMetAlaGlu 400
DB 1143 CGACGA----- 1148
QY 401 AlaTyrSerGluIleGlyThrLysGlyLysLysArgArgGlyLys 415
DB 1149 -----GGCGCGCGCGGCAACCGCAGCGGTGAG 1178

RESULT 2
US-08-465-473B-13
Sequence 13, Application US/08465473B
Patent No. 5939531
GENERAL INFORMATION:
APPLICANT: Weib, Winfried S.
APPLICANT: Hynes, Nancy E.
APPLICANT: Harwerth, Ina-Maria
APPLICANT: Groner, Bernd
APPLICANT: Hardman, No. 5939531man
APPLICANT: Zwickl, Markus
TITLE OF INVENTION: Recombinant Antibodies Specific for a
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: NOVARTIS Corporation
STREET: 564 Morris Avenue
CITY: Summit
STATE: New Jersey
COUNTRY: USA
ZIP: 07901-6940
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,473B
FILING DATE: 5 June 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/828,832
FILING DATE: 31-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 91-810079.3
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Pfeiffer, Henna J.
REGISTRATION NUMBER: 22,640
REFERENCE/DOCKET NUMBER: 4-18518/A/CIP/CONT2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908)522,6940
TELEFAX: (908)522,6955
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 2012 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
HYPOTHEICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Mouse/Pseudomonas aeruginosa
INDIVIDUAL ISOLATE: E. coli
IMMEDIATE SOURCE:
CLONE: pmw215-5
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1911 /note= "64-87 FLAG peptide and
OTHER INFORMATION: enterokinase cleavage site; 97-453 heavy chain
OTHER INFORMATION: variable domain;454-498 15 aa linker sequence;
OTHER INFORMATION: 499-822 FRP5 light chain variable domain
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 1..63
OTHER INFORMATION: /note= "ompA signal peptide"
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 94..1911
FEATURE:
NAME/KEY: 3'UTR

LOCATION: 1912..2012
OTHER INFORMATION: /function= "3'non-coding region of
OTHER INFORMATION: the exotoxin A gene"
FEATURE:
NAME/KEY: misc feature
LOCATION: 826..1911
OTHER INFORMATION: /note= "Exotoxin A gene coding
OTHER INFORMATION: region (coding for amino acids 252 to 613 of the
OTHER INFORMATION: mature exotoxin A)"
US-08-465-473B-13
Alignment Scores:
Pred. No.: 9,37e-117 Length: 2012
Score: 1310.50 Matches: 279
Percent Similarity: 68.51% Conservative: 19
Best Local Similarity: 64.14% Mismatches: 43
Query Match: 55.88% Gaps: 95
DB: 2 Index: 10
US-09-596-774-7 (1-443) X US-08-465-473B-13 (1-2012)
QY 1 GlnIleGlnLeuValGlnSerGlyProGlnLeuValysPProGlnIleThrValIle 20
DB 97 CAGGTACAACTGCAGCAGCTTGACCTGAACCTGAAGAAGCCTGAGAGACGTAAGATC 156
QY 21 SerCysValSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValysGlnAla 40
DB 157 TCTGCAAGGCTCTGGGTATCTTTCAACAACCTATGATGAATGAACTGGTGAACAGGCT 216
QY 41 ProGlnGlnGlyLeuValTyrPheGlyTyrIleAsnThrSerThrGlnIleSerThrPhe 60
DB 217 CCAGGACAGGGTTAAAGTGGATGGCTGATTAACACCTGCACCTGGAGAGCTCAACATTT 276
QY 61 AlaAspAspPheValysGlyArgPheAspPheSerIleuGlnThrSerAlaAsnThrAlaTyr 80
DB 277 GCTATATGATTCAGAGGAGCAGGTTTGACTCTCTTTGAAACCTGTGCACACCTGCTAT 336
QY 81 LeuGlnIleAsnLeuValysSerGlyAspMetAlaThrTyrPheCysAlaArgTyrGln 100
DB 337 TTGCAGATCAACAACCTCAAGAGTGAAGACATGCTCATATTTCTGTGCAAGATGGAG 396
QY 101 ValTyrHisGlyTyrValProTyrTyrGlnGlnIleThrValThrValSerSerGly 120
DB 397 GTTACCAACGGCTACGTTCTTACTGAGGGCCAAAGGACACAGGTCACCGTTCTCTGGC 456
QY 121 GlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlyGlySerAspIleGlnIleThrGln 140
DB 457 GGTGGCGGTCTGCTGGCGGTGGCTCCGCGGTGCGGTCTCTGACATCCAGCTACCCAG 516
QY 141 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysValAlaSer 160
DB 517 TCTCAAAATTCCTGTCCACTCACTGATGAGACAGGAGCATCATCTGCAAGGCCAGT 576
QY 161 GlnAspValTyrAsnAlaValAlaTyrTyrGlnGlnIleValysPProGlnIleSerProLysLeu 180
DB 577 CAGGATGTATATATGCTGTGGCTGGTATCAACAGAAACAGAGACATCTCTAAACTT 636
QY 181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200
DB 637 CTGATTTACTCGGATCTCCCGGTACACGAGATCTCTCTCTCCTCAGCGAGGAGTGC 696
QY 201 SerGlyProAspPheThrPheThrIleSerSerValGlnIleGlnAspLeuAlaValTyr 220
DB 697 TCTGGCGGATTTCACTTTCACATCAACAGATGAGGCTGAGGCTGAGGCTGAGGCTTAT 756
QY 221 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyTyrIleGlnIle 240
DB 757 TTCTGTCAAGCAACTTTTTCGATTCATTCACGTTCCGCTGGGAGCAAAATTCGATGC 816
QY 241 LysAlaLeuGlu-----IleSerAsnSerValMetTyrPheSerSerValValPro 257
DB 817 AAGCTCTAGAGGCGGAGCGGCTGGCCGCGCTGACCGGACCGAGCGCTGCACTGCGC 876

GENERAL INFORMATION:
APPLICANT: WELLS, Winfried, Dr.
APPLICANT: FOYMINAYA, Jesus
TITLE OF INVENTION: NUCLEIC ACID TRANSFER SYSTEM
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP
STREET: 655 15th St., N.W., Suite 330 - G St. Lobby
City: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-5701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/840,713
FILING DATE: 25-APR-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Kites, Monica Chin
REGISTRATION NUMBER: 36,105
REFERENCE/DOCKET NUMBER: 1614-7014
TELEPHONE: (202) 638 - 5000
TELEFAX: (202) 638 - 4810
INFORMATION FOR SEQ ID NO: 34:
SEQUENCE CHARACTERISTICS:
LENGTH: 1919 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
FEATURE:
NAME/KEY: CDS
LOCATION: 64..1908
US-08-840-713-34

Alignment Scores:
Pred. No.: 1,52e-116 Length: 1919
Score: 1308.00 Matches: 245
Percent Similarity: 97.64% Conservative: 3
Best Local Similarity: 96.46% Mismatches: 6
Query Match: 55.78% Indels: 0
Gaps: 0
DB: 4

US-09-596-774-7 (1-443) x US-08-840-713-34 (1-1919)

QY 1 Glnlleglnleuvalglnserglyprogluleuylsyserprogllyglnthrvallvle 20
DB 1147 CAGGTACACTGCAGCAGCTGCAGCTGAACTGAAAGAGCTGAGAGACAGTAAAGATC 1206
QY 21 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrPvalylsGlnAla 40
DB 1207 TCTGTGCAAGGCTTGGGTATCTCTTCAACAACATATGGAATGAATGGGTGAAGCAGGCT 1266
QY 41 ProglYglnlyleuLYserTyrMetGlyTyrIleAsnThrSerThrGlyGlnSerThrPhe 60
DB 1267 CCAGGACAGGGTTAAAGTGGATGGCTGATTAACACCTCACTGAGAGATCAACATTT 1326
QY 61 AlaAspAspPheLYserGlyAArgPheAspPheSerIleuGlnThrSerAlaAsnThrAlaTyr 80
DB 1327 GCTGATGACTTCAAGGAGCAGGCTTGTGACTTCTTGGAAACCTCTGCCAACACAGCTAT 1386
QY 81 LeuGlnIleAsnAsnLeuLYserGlyLysAspMetAlaThrTyrPheCysAlaArgTyrGln 100
DB 1387 TTGCAGATTCACACACTCAAAAGGAAGACATGCTCATATTTCTGTGCAAGATGGAG 1446
QY 101 ValTyrHisGlyTyrValProTyrTyrGlyGlnlyThrThrValThrValSerSerGly 120
DB 1447 GTTACACAGCGCTATGTTCTTACTGGGGCCAAAGGACACAGGTTCACCGTTCTCTGGC 1506

QY 121 GlyGlyGlySerGlyGlyGlySerGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 140
DB 1507 GTGGCGGCTTCTGTGGCGGTGGCTCCGGCGGTGGCGGTCTGACATCAGCTGACCCAG 1566
QY 141 SerHisLYserPheLeuSerThrSerValGlyAspArgValSerIleThrCysLYsAlaSer 160
DB 1567 TCTCACAAATTCCTGTCTCATTGAGTGAACAGAGGCTCAGATCCTCAAGGCGCAGT 1626
QY 161 GlnAspValTyrAsnAlaValAlaTyrTyrGlnlySerProglYglnserProLYleu 180
DB 1627 CAGATGTGATATGCTGTGCTGTATCAACAGAAACAGACATCTCTAAACTT 1686
QY 181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200
DB 1687 CTGATTTACTCGGATCTCCCGGTACATCGAGATCCCTTCTGCTTCACTGCGAGTGC 1746
QY 201 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 220
DB 1747 TCTGGCGCGATTTCATCTTCAACATCAGAGTGTGAGGCTGAAGACTGGCAGTTAT 1806
QY 221 PheCysGlnIleHisPheArgThrProPheThrPheGlySerGlyThrLYsLeuGlnIle 240
DB 1807 TTCTGTACGACAACTTTCTGACTCATTGACGTTGGCTCGGGGACAAATTTGAGATC 1866
QY 241 LysAlaLeuGlnIleSerAsnSerValMetTyrPheSerSer 254
DB 1867 AAGCTTACAGATCTCTGAGTGAAGAAATTTTACCC 1908

RESULT 5
US-08-840-713-1
Sequence 1, Application US/08840713
Patent No. 6498233
GENERAL INFORMATION:
APPLICANT: WELLS, Winfried, Dr.
APPLICANT: FOYMINAYA, Jesus
TITLE OF INVENTION: NUCLEIC ACID TRANSFER SYSTEM
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP
STREET: 655 15th St., N.W., Suite 330 - G St. Lobby
City: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-5701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/840,713
FILING DATE: 25-APR-1997
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Kites, Monica Chin
REGISTRATION NUMBER: 36,105
REFERENCE/DOCKET NUMBER: 1614-7014
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 638 - 5000
TELEFAX: (202) 638 - 4810
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1692 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
IMMEDIATE SOURCE:
CLONE: pmf46-5
FEATURE:
NAME/KEY: sig peptide
LOCATION: 1..63

OTHER INFORMATION: /product= "E. coli ompa signal"
 OTHER INFORMATION: peptide"
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 64..1656
 OTHER INFORMATION: /product= "scfV (FRP5) -delta"
 OTHER INFORMATION: ETA-delta GAL4"
 US-08-840-713-1

Alignment Scores:

Pred. No.:	2.47e-116	Length:	1692
Score:	1305.00	Matches:	278
Percent Similarity:	68.36%	Conservative:	18
Best Local Similarity:	64.20%	Mismatches:	43
Query Match:	55.65%	Indels:	95
DB:	4	Gaps:	10

US-09-596-774-7 (1-443) x US-08-840-713-1 (1-1692)

```

QY 1 GlnIleGlnLeuValGlnSerGlyProGluLeuValysPProGlyGluThrValysIle 20
DB 115 CAGGTACACGCGACGAGCTGAGACTGAAAGAGCCTGGAGAGACAGTCAAGATC 174
QY 21 SerGlyValAspSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValysGlnAla 40
DB 175 TCCGCAAGGCGCTCGGGTATCTTTCACAACTATGGAATGAATGAACTGGGTGAAAGAGGCT 234
QY 41 ProGlyGlnGlyLeuValysTyrPheSerGlyTyrIleAsnThrSerThrGlyGluSerThrPhe 60
DB 235 CCAGGACAGGCGTTAAAGTGAATGAGGCTGAGTTACACTTCACTGAGAGTCAACATTT 294
QY 61 AlaAspAspPheValysGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80
DB 295 GCTGATGACTTCAAGGAGCGGTTGACTTCTTGGAAACCTTGCACACCTGCTAT 354
QY 81 LeuGlnIleAsnLeuValysSerGlyValAspMetAlaThrTyrPheCysAlaArgTyrGlu 100
DB 355 TTGAGATCAACACCTCAAAAGAAAGACATGCTCATTTCTGTGCAAGATGGAG 414
QY 101 ValTyrAsnGlyTyrValProTyrTyrGlnGlyThrThrValThrValSerSerGly 120
DB 415 GTTTCACACGCGCTACGTTCTTCTTCTGAGGCGCAAGGACACGACCTTCTCTGGC 474
QY 121 GlyGlyValysSerGlyGlyGlySerGlyGlyGlyValysSerAspIleGlnLeuThrGln 140
DB 475 GGTGGCGGTTCTGGTGGCGGTGGCTCCGCGGTGGCTTGTGACATCCAGCTGACCCAG 534
QY 141 SerHisValPheLeuSerThrSerValGlyAspArgValSerIleThrCysValAspSer 160
DB 535 TCTCACAATTCCTGTCACCTCAGTGAAGAGACAGGCTCAGCATCAGCTCAGAGCCAGT 594
QY 161 GlnAspValTyrAsnAlaValAlaTyrTyrGlnGlnIleAspProGlyGlnSerProLysLeu 180
DB 595 CAGAGTGTGTAATGCTGCTGCTGCTGATCAACAGAAACAGACATCTCTTAACATT 654
QY 181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200
DB 655 CTGATTACTCGGATCTCCCGGTACACAGAGTCCCTTCTCGCTTCACTGGAGAGGC 714
QY 201 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 220
DB 715 TCTGGGCGGATTCCTTTCACATCAGAGTGAAGTGAAGCTGAAAGCTGGAGATTAT 774
QY 221 PheCysGlnGlnHisPheArgTyrProPheThrPheGlySerGlyThrLysLeuGluIle 240
DB 775 TTCTGTAGAGCACTTTCTTCTTCTTCACTTCACTTCCGCTCGGAGCAAAATTCGAGATC 834
QY 241 LysAlaLeuGlu-----IleSerAsnSerValMetTyrPheSerSerValValPro 257
DB 835 AAAGCTCTAGAGCGCGACGCGCTGACCGCTGACCGACGACGAGGCTCCACCTGGCG 894
QY 258 ValLeuGlnIleValAsnSerThrThrThrLys----- 268
  
```

```

DB 895 CTG-----GAGACTTTCACCCGTCATCGCAGCGCGCTGGAAACA 939
QY 269 -----ProValLeuArgThrProSerProValHisProThrGlyThr 282
DB 940 CTGAGCAGATGCGCGCTATCCGTCGACGCGCTGCT- GCGCTCTTA---CCTGGCGGCG 995
QY 283 Ser-----GlnProGlnArgProGluAspCysArgProArgIleValysGlyThr 300
DB 996 ACTGTATGGAACCAAGTTCACACAGATGATCCGCAACGCCCTGGCCAGCCCCGAGCGG 1055
QY 301 GlyLeuAspPheLeuGluAspProLysLeuCysTyrLeuLeuAspGlyIleLeuPheIle 320
DB 1056 CGGCGACCTGGCGAAGCGATCCG----- 1079
QY 321 TyrGlyValIleThrAlaLeuTyrLeuArgAlaLysPheSerArgSerAlaGluThr 340
DB 1080 -----CGAGCAGCCGAGACA 1094
QY 341 AlaAlaAsnLeuGlnAspProAsnGlnLeuTyrAsnGlnLeuAsnLeuGlyArgArgGlu 360
DB 1095 GCGCCGCTGCGCCCTGACCT-----GCGCCGCCCA 1127
QY 361 GluTyrAspValLeuGluLysValArgAlaArgAspProGluMetGlyLysGlnGln 380
DB 1128 GAG-----CGAGCGCTTCGTCCG-----CGAGGCGCACCGCAA 1160
QY 381 ArgArgArgAsnProGlnGluGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaGlu 400
DB 1161 CGACGA----- 1166
QY 401 AlaTyrSerGluIleGlyThrLysGlyGlyArgArgArg 413
DB 1167 -----GCGCGCGCGCGCAACGCGACA 1190
  
```

RESULT 6

US-08-235-838-6
 Sequence 6, Application US/08235838
 Patent No. 5571894

GENERAL INFORMATION:

APPLICANT: Wells, Winfried S.
 APPLICANT: Hynes, Nancy E.
 APPLICANT: Harwerth, Ina-Maria
 APPLICANT: Groner, Bernd
 APPLICANT: Hardman, No. 5571894man
 TITLE OF INVENTION: Recombinant Antibodies Specific for a
 TITLE OF INVENTION: Growth Factor Receptor
 NUMBER OF SEQUENCES: 16
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: CIBA-GEIGY Corporation
 STREET: 7 Skyline Drive
 CITY: Hawthorne
 STATE: New York
 COUNTRY: USA
 ZIP: 10532
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/235,838
 FILING DATE: TBA
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 07/828,832
 FILING DATE: 31-JAN-1992
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: GB 91-810079.3
 FILING DATE: 05-FEB-1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Elmer, James Scott
 REGISTRATION NUMBER: 36,129

REFERENCE/DOCKET NUMBER: 4-18518/A/CIP/CONT
TELECOMMUNICATION INFORMATION:
TELEPHONE: (919) 541-8614
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 2233 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Mouse and E. coli
INDIVIDUAL ISOLATE: E. coli
IMMEDIATE SOURCE:
CLONE: pMW616
FEATURE:
NAME/KEY: CDS
LOCATION: 23..2155
OTHER INFORMATION: /note= "89-445 FRP5 heavy chain
var. domain; 446-490 is aa linker sequence; 491-814
OTHER INFORMATION: FRP5 light chain var. domain; 815-2155 coding region
OTHER INFORMATION: of phoA
FEATURE:
NAME/KEY: 5' UTR
LOCATION: 1..22
OTHER INFORMATION: /function= "ompa 5' non-coding
OTHER INFORMATION: region"
FEATURE:
NAME/KEY: sig peptide
LOCATION: 23..85
OTHER INFORMATION: /note= "ompa signal peptide"
FEATURE:
NAME/KEY: 3' UTR
LOCATION: 2156..2233
OTHER INFORMATION: /function= "phoA 3' non-coding
OTHER INFORMATION: region"
FEATURE:
NAME/KEY: mat peptide
LOCATION: 86..2155
US-08-235-838-6
Alignment Scores:
Pred. No.: 5.8e-116 Length: 2233
Score: 1303.00 Matches: 245
Percent Similarity: 95.00% Conservative: 2
Best Local Similarity: 94.23% Mismatches: 1
Query Match: 55.57% Indels: 12
DB: 1 Gaps: 1
US-09-596-774-7 (1-443) x US-08-235-838-6 (1-2233)
QY 1 Gln1leglnleuValGlnserGlyProgluLeuLysLysProgluLysLysLeuThrValLysLeu 20
DB 89 CAGGTACCACTGCGACGCTGACCTGACCTGAAGAGCGCTGGAGACAGTCAGATC 148
QY 21 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValLysGlnAla 40
DB 149 TCCGCAAGGCTCTGGGTATCCTTTCACAACTATGGAATGAATGAGTGAACAGGCT 208
QY 41 ProgluLysLysLeuLysLysLysLysLysLysLysLysLysLysLysLysLysLysLys 60
DB 209 CCAGACAGGCTTTAAAGTGGATGGCTGAGTTAACACCTCCAGCTGAGAGTCAACATTT 268
QY 61 AlaAspAspPheLysGlyLysPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80
DB 269 GCTATGACTTCAAGGAGCGGTTGACTTCTTTGGAAACCTCTGCAACAGCTGCTAT 328
QY 81 LeuGln1leuAsnLeuLysSerGlyLysPheAlaThrTyrPheCysAlaLysLysLysLys 100
DB 329 TTGCAATCAACCACTCAAAAGTGAAGACATGGCTATATTTCTGTGCAAGATGGGAG 388

QY 101 ValTyrHisGlyTyrValProTyrTyrGlyGlnGlyThrThrValThrValSerSerGly 120
DB 389 GTTACCAACGCGTCTGCTTCTTACTGGGCGCAAGGACACAGGTCACGTTCTCTGGC 448
QY 121 GlyGlySerGlyGlyGlyGlySerGlyGlyGlyGlySerAsp1leglnLeuThrGln 140
DB 449 GGTGGCGTTCTGGTGGCGGTTGGCTCCGGCGGTGGCGGTCTGACATTCAGCTACCCAG 508
QY 141 SerHisLysPheLeuSerThrSerValGlyAspArgAlaSer1leuThrCysAlaLys 160
DB 509 TCTCACAAAATTCCTGTCACCTCAGTAGAGACAGGCTCAGCATCCTGCAAGCCAGT 568
QY 161 GlnAspValTyrAsnAlaValAlaTyrTyrGlnGlnLysProgluLysSerProLysLeu 180
DB 569 CAGATGTGTAATATGCTGTGGCTGTATCAACAGAAACAGACAAATCTCTAACTT 628
QY 181 Leu1leuTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200
DB 629 CTGATTTACTCGGCATCTCCCGGTACACTGAGAGTCCCTTCTGCTTCACTGGCAGTGC 688
QY 201 SerGlyProAspPheThrPheThrLysSerSerValGlnAlaGluAspLeuAlaValTyr 220
DB 689 TCTGGCGCGATTTCTTCTGATCTCCTTCACTGCTGGGCGGCAAAATTCAGATC 748
QY 221 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyTyrLysLeuGluLe 240
DB 749 TTCTGTACGACCACTTTTCTGATCTCCTTCACTGCTGGGCGGCAAAATTCAGATC 808
QY 241 LysAlaLeuGluLysSerAsnSerValMetTyrPheSerSerValProValLeuGln 260
DB 809 AAAGCTTAGG-----CTGTTCTGGAA 832
RESULT 7
US-08-465-473B-6
Sequence 6, Application US/08465473B
Patent No. 5939531
GENERAL INFORMATION:
APPLICANT: Wels, Winfried S.
APPLICANT: Hynes, Nancy B.
APPLICANT: Harwerth, Ina-Maria
APPLICANT: Groner, Bernd
APPLICANT: Hardman, No. 5939531man
APPLICANT: Zwickl, Markus
TITLE OF INVENTION: Recombinant Antibodies Specific for a
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: NOVARTIS Corporation
STREET: 564 Morris Avenue
CITY: Summit
STATE: New Jersey
COUNTRY: USA
ZIP: 07901-6940
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,473B
FILING DATE: 5 June 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/828,832
FILING DATE: 31-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 91-810079.3
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Pfeiffer, Henna J.
REGISTRATION NUMBER: 22,640
REFERENCE/DOCKET NUMBER: 4-18518/A/CIP/CONT2

TELECOMMUNICATION INFORMATION:
TELEPHONE: (908)522.6940
TELEFAX: (908)522.6955
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 2233 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Mouse and E. coli
INDIVIDUAL ISOLATE: E. coli
IMMEDIATE SOURCE:
CLONE: pMW616
FEATURE:
NAME/KEY: CDS
LOCATION: 23..2155 /note= "89-445 FRP5 heavy chain
OTHER INFORMATION: var.domain: 446-490 15 aa linker sequence: 491-814
OTHER INFORMATION: FRP5 light chain var.domain: 815-2155 coding region
OTHER INFORMATION: of phoA
FEATURE:
NAME/KEY: 5'UTR
LOCATION: 1..22
OTHER INFORMATION: /function= "ompA 5'non-coding
OTHER INFORMATION: region"
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 23..85
OTHER INFORMATION: /note= "ompA signal peptide"
FEATURE:
NAME/KEY: 3'UTR
LOCATION: 2156..2233
OTHER INFORMATION: /function= "phoA 3' non-coding
OTHER INFORMATION: region"
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 86..2155
US-08-465-473B-6
Alignment Scores:
Pred. No.: 5.8e-116 Length: 2233
Score: 1303.00 Matches: 245
Percent Similarity: 95.00% Conservative: 2
Best Local Similarity: 94.23% Mismatches: 1
Query Match: 55.57% Indels: 12
DB: 2 Gaps: 1
US-09-596-774-7 (1-443) x US-08-465-473B-6 (1-2233)
QY 1 GlnllleGlnleuValGlnSerGlyProGluLeuLysProGlyGluThrValLysIle 20
DB 89 CAGGTAACTGACGACGTGTGACCTGAACTGAAGAGCGCTGGAGAGACAGTCAGATC 148
QY 21 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTrpValLysGlnAla 40
DB 149 TCCTGCAGAGCGCTGTGGTATCTCTTACAAACTATGAAATGAATGGGTGGAGGCAAGCT 208
QY 41 ProGlyGlnGlyLeuLysTyrMetGlyTyrIleAsnThrSerThrglyGluSerThrPhe 60
DB 209 CCAGGACAGGCGTTAAAGTGATGGGCTGGATTAAACCTCCACCTGGAGAGTCACAQTTT 268
QY 61 AlaAspAspPheLysGlyTyrGlnPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80
DB 269 GCTGATGACTTCAAGGAGCGGTTGACTTCTTGGAAACCTCTGCAACACCTGCCTAT 328
QY 81 LeuGlnIleAsnLysLeuLysSerGlyAspMetAlaThrTyrPheCysAlaTGTpGlu 100
DB 329 TTGCAGATCAACACCTCAAAAGTGAAGACATGGCTACATATTCTGTGCAAGATGGAG 388

QY 101 ValTyrHisGlyTyrValProTyrTrpGlyGlnGlyThrThrValThrValSerSerGly 120
DB 389 GTTACACAGCGCTACGTTCTTACTGGGACCAAGGAGCAAGGTCACCGTTCTCTCGGC 448
QY 121 GlyGlyGlySerGlyGlyGlySerGlyGlyGlyGlyGlySerAspIleGlnLeuThrGln 140
DB 449 GGTGGCGGTTCTGTGGCGGTGGGCTCCGAGCGGTGGGCTTCAACATCCAGCTGACCCAG 508
QY 141 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSer 160
DB 509 TCTCACAATTCCTGTCCTCCTCAGTAGACAGGAGTCCAGTACCTCCAGAGGCGCAGT 568
QY 161 GlAspValTyrAsnAlaValAlaTyrGlnGlnLysProGlyGlnSerProLysLeu 180
DB 569 CAGATGTGTATATGCTGTGCTGGTATCAACAGAACAGACAGACATCTCTTAACCTT 628
QY 181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200
DB 629 CTGATTTACTCGGCATCTCCCGGTACACTGAGATCCCTTCTGCTTCACTGGCAGTGGC 688
QY 201 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 220
DB 689 TCTGGCGCGATTTCACCTTCAACATCAACAGTGTGACAGGCTGMAAGCTGGCAGTTTAT 748
QY 221 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyThrLysLeuGluIle 240
DB 749 TTCTGTACGAAACATTTTCTGATCTTCACTTCACTTGGCTGGGAGCAAAATTGGAGATC 808
QY 241 LysAlaLeuGluIleSerAsnSerValMetTyrPheSerSerValValProValLeuGln 260
DB 809 AAGCTCTAGAG-----CTGTGTCGAA 832

RESULT 8
US-08-235-838-4
Sequence 4, Application US/08235838
Patent No. 5571894
GENERAL INFORMATION:
APPLICANT: Weis, Winfried S.
APPLICANT: Hynes, Nancy E.
APPLICANT: Harwerth, Ina-Maria
APPLICANT: Groner, Bernd
APPLICANT: Hardman, No. 5571894man
APPLICANT: Zwickl, Markus
TITLE OF INVENTION: Recombinant Antibodies Specific for a
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: CIBA-GEIGY Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: New York
COUNTRY: USA
ZIP: 10512
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/235,838
FILING DATE: TBA
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/828,832
FILING DATE: 31-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 91-810079.3
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET NUMBER: 4-18518/A/CIP/CONT
TELECOMMUNICATION INFORMATION:

TELEPHONE: (919) 541-8614
TELEFAX: (919) 541-8689
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 748 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
HYPOTHEICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE: Mouse
INDIVIDUAL ISOLATE: E. coli
IMMEDIATE SOURCE:
CLONE: pMW52
FEATURE:
NAME/KEY: CDS
LOCATION: 6..728
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..701 /note= "1-8 synthetic spacer;9-365
OTHER INFORMATION: FRP5 heavy chain var. domain;99-113 CDR1H;156-206
OTHER INFORMATION: CDR2H;303-332 CDR3H;366-410 15 aa linker seq;411-728
OTHER INFORMATION: FRP5 light chain var dom;480-512 CDR1L;558-578 CDR2L;
OTHER INFORMATION: 675-701 CDR3L
US-08-235-838-4

Alignment Scores:
Pred. No.: 5,76-115 Length: 748
Score: 1285.50 Matches: 241
Percent Similarity: 98.37% Conservative: 1
Best Local Similarity: 97.97% Mismatches: 1
Query Match: 54.82% Indels: 3
DB: 1 Gaps: 1

US-09-596-774-7 (1-443) x US-08-235-838-4 (1-748)

QY 1 Glnlleglnleuvalglnsergylprogluleuylsyproglygluthvallygllle 20
DB 9 CAGGTACAACTGCAGCAGTCTGCACTGAACTGAAGAACTGGAAGACAGTCAAGATC 68
QY 21 SerCysValysAlaserglytyrProPheThraentyrGlyMeaanttrpvallyglnala 40
DB 69 TCCGTGCAAGGCTTGGGTATCTCTTTCACAACTATGGAAGAACTGGGTGAAGCAGGCT 128
QY 41 Proglyngllyleuylstrpmetgyltrpilleasnthrserthnglygluserthrphe 60
DB 129 CCAGGACAGGCTTAAAGTGAATGGGCTGATTAACATTCACATGGAAGATCAACATTT 188
QY 61 AlaAspPhePheylsrglytyrGlyPheAspPheSerleugluThrSerAlaThrAlaTyr 80
DB 189 GCTGATGACTTCAAGGAGCGGTGACTTCTCTTGGAACTCTGCCAACACTGCGCTAT 248
QY 81 LeuglnlleasnaentleuylsersergluspmeclaThrtyrPheCysAlaArgTrpGlu 100
DB 249 TTGAGAGATCAACAACCTCAAAAGTGAAGACATGGCTACATATTTCTGTGCAAGATGGAG 308
QY 101 ValtyrHlsglytyrvalprolytrpqllynglylthrrhvalthrvalsersergly 120
DB 309 GTTACACAGGCTACGTTCTTACTGGGGCCAGAGGACCAAGGTACCGTTCTCTGGC 368
QY 121 Glyglyserglyglyglyglyserglyglyglyseraplleglneuthrgln 140
DB 369 GGTCGCGGTTCTGGTGGGCGGTGCTCCGCGGTGCTTGCATCCACCTGACCCAG 428
QY 141 SerHisLysPheLuserThrservalglyAspargvalserIethrCysAlaLser 160
DB 429 TCTCAAAATTCCTGCTCAGTACAGAGACAGGAGTCAACATCCTGCAAGGCGCT 488
QY 161 GlnAspValtyrAsnaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 180

DB 489 CAGGATGTGATTAATGCTGTGCTGGTATCACAAGAAACAGACAAATCTCTAACTT 548
QY 181 LeuIleTySerAlaSerArglytyrThrGlyValProSerArgPheThrGlySerGly 200
DB 549 CTGATTTACTCGGATCTCTCCGGTATACATGAGATCCCTTCTCCCTTCACTGGCAGTGGC 608
QY 201 SerGlyProAspPheThrPheThrIleSerSerValglnAlaGluAspLeuAlaValtyr 220
DB 609 TCTGGCCGAGTTTCACTTTCACATCAGAGTGTGCAAGGCTGAAGACCTGGCAGTTTAT 668
QY 221 PheCysGlnGlnHisPheAlaGlyThrProPheThrPheGlySerGlytyrThrlyleuGlu--- 239
DB 669 TTCTGTGAGCAAACTTTTCGTACTCAATCAACGTTCCGCTCGGGAGCAAAATGGAGATC 728
QY 240 -----IleLysAlaLeu 243
DB 729 TAGCTGATCAAGCTCTTA 746
RESULT 9
US-08-465-473B-4
Sequence 4, Application US/08465473B
Patent No. 5939531
GENERAL INFORMATION:
APPLICANT: Wele, Winfried S.
APPLICANT: Hynes, Nancy E.
APPLICANT: Harwerth, Ina-Maria
APPLICANT: Groner, Bernd
APPLICANT: Hardman, No. 5939531man
APPLICANT: Zwickl, Markus
TITLE OF INVENTION: Recombinant Antibodies Specific for a
TITLE OF INVENTION: Growth Factor Receptor
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: NOVARTIS Corporation
STREET: 564 Morris Avenue
CITY: Summit
STATE: New Jersey
COUNTRY: USA
ZIP: 07901-6940
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,473B
FILING DATE: 5 June 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/828,832
FILING DATE: 31-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 91-810079.3
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Pfeiffer, Henna J.
REGISTRATION NUMBER: 22,640
REFERENCE/DOCKET NUMBER: 4-18518/A/CIP/CONT2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908) 522 6940
TELEFAX: (908) 522 6955
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 748 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
HYPOTHEICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE: Mouse
INDIVIDUAL ISOLATE: E. coli

IMMEDIATE SOURCE:
CLONE: pMW52
FEATURE:
NAME/KEY: CDS
LOCATION: 6..728
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..701
OTHER INFORMATION: /note="1-8 synthetic spacer;9-365
OTHER INFORMATION: FRP5 heavy chain var domain;99-113 CDRH;156-206
OTHER INFORMATION: CDRH;303-332 CDRH;366-410 15 aa linker seq;411-728
OTHER INFORMATION: FRP5 light chain var dom;480-512 CDRH;558-578 CDR2L;
OTHER INFORMATION: 675-701 CDR3L
US-08-465-473B-4

Alignment Scores:

Pred. No.: 5.7e-115 Length: 748
Score: 1285.50 Matches: 241
Percent Similarity: 98.37% Conservative: 1
Best Local Similarity: 97.97% Mismatches: 1
Query Match: 54.82% Indels: 3
DB: Gaps: 1

US-09-596-774-7 (1-443) x US-08-465-473B-4 (1-748)

QY 1 GlnIleGlnLeuValGlnSerGlyProGlyLeuLysProGlyGlyLysIle 20
DB 9 CAGGTACACAGCGACAGCTGACCTGACCTGAAGAGAGAGAGAGAGAGATC 68
QY 21 SerCysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTPValysGlnAla 40
DB 69 TCCTGCAAGGCGCTCTGGGATCTCTTCAAACTATGGAATGAACCTGGGAGAGAGGCT 128
QY 41 ProGlyGlnGlyLeuLysTPMetGlyTyrIleAsnThrSerThrGlyLysThrPhe 60
DB 129 CCAAGGAGAGGCTTTAAAGTGGATGGCTGGATTAACCTTCCACTGAGAGATCAACATTT 188
QY 61 AlaAspAspPheLysGlyTyrPheAspPheSerLeuGlyThrSerAlaAsnThrAlaTyr 80
DB 189 GCTGATACCTTCAAGGAGCGGTTGACTTCTTGGAAACCTCTGCAACAGCGCTAT 248
QY 81 LeuGlnIleAsnLeuLysSerGlyAspMetAlaThrTyrPheCysAlaArgTPGlu 100
DB 249 TTGAGATTCACACACCTCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGATGGAG 308
QY 101 ValTyrHisGlyTyrValProTyrTPGlyGlnGlyThrValThrValSerSerGly 120
DB 309 GTTTACCAAGGCTACGTTCTTACTGGGAGGAGAGAGAGAGAGAGAGAGAGAGAG 368
QY 121 GlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThrGln 140
DB 369 GGTGGCGGCTCTGGTGGCGGAGGCTCCGGGCGGCTTCTGACATCCAGCAGCCAG 428
QY 141 SerHisAlaPheLeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSer 160
DB 429 TCTCACAAATTCCTGTCCTCAGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 488
QY 161 GluAspValTyrAsnAlaValAlaTyrGlnGlyLysProGlyGlnSerProLysLeu 180
DB 489 CAGAGTGTGTAATGCTGTGCTGGATCAACAGAAACAGAGACATCTCTAAACCTT 548
QY 181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200
DB 549 CTGATTACTCGGATCTCCCGGTACACTGAGAGCTCTTCTGCTTCACTGGCAGTGGC 608
QY 201 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 220
DB 609 TCTGGGCGGATTTCACTTTCACATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 668
QY 221 PheCysGlnGlnHisPheArgTyrProPheThrPheGlySerGlyThrLysLeuGlu--- 239
DB 669 TTCTGTACGACAACTTTTCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTT 728

QY 240 -----IleLysAlaLeu 243
DB 729 TAGCTGATCAAGCTCTTA 746

RESULT 10

US-08-356-786-15

Sequence 15, Application US/08356786

Patent No. 5877305

GENERAL INFORMATION:

APPLICANT: Huston, James S.

APPLICANT: Oppermann, Hermann

APPLICANT: Houston, L. L.

APPLICANT: Ring, David B.

TITLE OF INVENTION: Biosynthetic Binding Protein for Cancer

TITLE OF INVENTION: Marker

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: Edmund R. Pitcher, Testa, Hurwitz, & Thibault

STREET: Exchange Place, 53 State Street

City: Boston

STATE: Massachusetts

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/356,786

FILING DATE:

CLASSIFICATION: 424

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/831,967

FILING DATE: 06-FEB-1992

ATTORNEY/AGENT INFORMATION:

NAME: Pitcher, Edmund R.

REGISTRATION NUMBER: 27,829

REFERENCE/DOCKET NUMBER: CRP-053

TELEPHONE: (617) 248-7000

TELEFAX: (617) 248-7100

INFORMATION FOR SEQ ID NO: 15:

SEQUENCE CHARACTERISTICS:

LENGTH: 1869 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)

FEATURE:

NAME/KEY: CDS

LOCATION: 1..1869

OTHER INFORMATION: /note="Product = "7418Fv-PB40"

US-08-356-786-15

Alignment Scores:

Pred. No.: 2.72e-85 Length: 1869
Score: 984.50 Matches: 226
Percent Similarity: 61.52% Conservative: 41
Best Local Similarity: 52.07% Mismatches: 90
Query Match: 41.98% Indels: 78
DB: Gaps: 12

US-09-596-774-7 (1-443) x US-08-356-786-15 (1-1869)

QY 1 GlnIleGlnLeuValGlnSerGlyProGlyLeuLysProGlyGlyLysIle 20
DB 7 GAGATTCATATGCTGACAGCTGAGCTGAAGAGAGCTGAGAGAGAGATC 66
QY 21 SerCysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTPValysGlnAla 40
DB 67 TCCTGCAAGGCTTCTGGGATTAACCTTCAAACTATGGAATGAACCTGGGAGAGGCT 126


```

Qy 41 ProGlyGlnGlyLeuLeuYsrTrpMetGlyYrTrpLeuSerThrGlyGlySerThrPhe 60
Db 127 CCAGGAAAGGCTTTAAAGTGAATGGCTGATTAACCAACCACTGAGAGCCAACTAT 186
Qy 61 AlaAspAspPheUyGlyArgPheAspPheSerLeuGlnThrSerAlaSerThrAlaTy 80
Db 187 GCTGAAGAGTTCAAGGAGCGGTTTGCTTCTTCTTGGAAACCTCTGCACAGCACTGCTAT 246
Qy 81 LeuGlnLeuSerLeuLeuYsrSerGlyLeuAspMetAlaThrYrPheCysAlaArgTrpGlu 100
Db 247 TTGGAGATTCACCAACCTCAAAAAGAGACAGGCTCACTATTTCTGTGAAGGCAATTT 306
Qy 101 ValTyriGlyYrValProTyTrpGlyGlnGlyThrThrValThrValSerSerGly 120
Db 307 ATTACTACAGCGGGGTTTGTACTGGGGCCAGAGGACTCTGTGCACTGTCTTCCGA--- 363
Qy 121 GlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlySerAlaPheLeuThrGln 140
Db 364 TCGAGCTCCCTCCGATCTTCATCTAGCGGTTCCAGCTGAGCGATATGTCATGACCCAG 423
Qy 141 SerHisLeuPheLeuSerThrSerValGlyAspArgValSerIleThrCysLeuAlaSer 160
Db 424 TCTCTTAATTCATGATCCATCGCTCAGTGGAGACAGGCTCAGCATCTCTCCAAAGCCAGT 483
Qy 161 GluAspValTyraAspAlaValAlaTrpTyriGlnGlnLeuYsrProGlyGlnSerProLeu 180
Db 484 CAGAGTGTGATGACTGCTGTAGCTGTGATATCAAAAACAGGAGCAATCTCTTAACTA 543
Qy 181 LeuIleTyriSerAlaSerSerArgTyriThrGlyValProSerArgPheThrGlySerGly 200
Db 544 CTGATTATCAGGACATCCACCCGACACCTGAGTCCCTGATCCGTTCAACGAGCAGTGA 603
Qy 201 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTy 220
Db 604 TCTGGGACAGATTTACTCTTCATCATCATCAGCATGCTGAGGCTGAAAGCTGGACCTCAT 663
Qy 221 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyThrLeuGlnIle 240
Db 664 TACTGTGAGCAACATTTAGAGTGCGCTTACCGTTCGAAAGGGAGCCAACTGGAATTA 723
Qy 241 LeuAlaLeuGlnIleSerAsnSerValMetYrPheSerSerValProValLeuGln 260
Db 724 AAACGGGCTGATGCTGCACCAACGTGA-----TCCATCTTCCCAACATCCAGT 771
Qy 261 LeuValAlaSerThrThrThrLysProValLeuArgThrProSerProValHis---Pro 279
Db 772 GAGCAGTTTGAGGGCGGAGCTGCGCCGCTGAACGCGCACAGGCTTGGCACCCTGCGC 831
Qy 280 ThrGlyThrSerGlnProGlnArgProGluAspCysArgProArgGlySerValLysGly 299
Db 832 CTGGAGACTTTACCCGTCATGCG-----CAGCCGCGCGC----- 867
Qy 300 ThrGlyLeuAspPheLeuGluAspProLysLeuCys---TyriLeuLeuAspGlyIleLeu 318
Db 868 -----TGGAAACAACGTGAGCAGCGCGCTATCCGCTGTCACAGCGCTGCTC 912
Qy 319 PheIleTyriGlyValIleIleThrAlaLeuTyriLeuArgAlaLysPheSer----- 335
Db 913 -----GCCCTTACCTGCGCGCGCGCTGCTGCTGGAACACAG 948
Qy 336 -----ArgSerAlaGluThrAla----- 341
Db 949 GTCGACCAAGTGATCCGCAACGCTCGGCGAGCCCGGCGAGCGGCGGACCTTGGCGAA 1008
Qy 342 AlaAsnLeuGlnAspProAsnGlnLeuTyraAsnGlnLeuAsnLeuGlyArgArgGlnGlu 361
Db 1009 GCGATTCGCGAGCAGCGCGAGCAGCGCGCTGCGCTGAGCCCT- GCGCGCGCGCAGAG 1067
Qy 362 TyrAspValLeuGlnLysLysArgAlaArgAspProGluMetGlyGlyLysGlnGlnArg 381
Db 1068 -----CGAGCGCTTGCCTCG-----GCAAGGCGCAACGCGCAACGA 1100
Qy 382 ArgArgAsnProGlnGlnGlyValTyriAspAlaLeuGlnLysAspLysMetAlaGlnAla 401

```

```

Db 1101 CGA----- 1103
Qy 402 TyriSerGlnIleGlyThrLysGlyGlnArgArgGlyLys 415
Db 1104 -----GGCGCGCGCGCAACGCGCGAGCTGCTGAG 1133

RESULT 11
US-08-133-804-1
; Sequence 1, Application US/08133804
; Patent No. 5534254
; GENERAL INFORMATION:
; APPLICANT: Huston, James S.
; APPLICANT: Oppermann, Hermann
; APPLICANT: Houston, L. L.
; APPLICANT: Ring, David B.
; TITLE OF INVENTION: Biosynthetic Binding Proteins For
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Testa, Hurwitz & Thibault/Patent Department
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/133,804
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Kelley, Robin D.
; REGISTRATION NUMBER: 34,637
; REFERENCE/DOCKET NUMBER: 2054/22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-248-7477
; TELEFAX: 617-248-7100
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 909 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..752
; OTHER INFORMATION: /product= "741F8 sFv' C-terminal
; OTHER INFORMATION: Gly4-Cys"
US-08-133-804-1

Alignment Scores:
Score: 1.84e-85 Length: 909
Percent Similarity: 85.89% Matches: 185
Best Local Similarity: 76.76% Conservative: 22
Query Match: 41.86% Mismatches: 33
DB: 1 Indels: 1 Gaps: 1

US-09-596-774-7 (1-443) x US-08-133-804-1 (1-909)
Qy 1 GlnIleGlnLeuValGlnSerGlyProGluLeuLysLysProGlyGlnThrValLysIle 20
Db 9 GAGATTCATTTGTCGAGCTTGACCTGAGCTGAAGAGCCTGAGAGACAGTCAAGATC 68
Qy 21 SerCysAlaAspSerGlyTyriProPheThrAsnTyriGlyMetAsnTrpValLysGlnAla 40
Db 69 TCTGCAAGGCTTCTGAGTATACCTTCAAACTATGGAATGAATGAGTGGTGAACAGGCT 128

```

```

QY      41  ProGlyGlnGlyLeuLeuYsrPmeGlyTrpIleAsnThrSerThnGlyGluSerThrPhe 60
Db      129  CCAGGAAGGCTTTAAAGTGGATGGCTGGATAAACCAACCACTGGAGAGCCACATAT 188
QY      61  AlaAspAspPheLeuGlyValArgPheAspPheSerLeuGlnThrSerAlaAsnThrAlaTyr 80
Db      189  GCTGAAGAGTTCAAGGAGCGGTTGGCTTCTCTTGGAAACCTCTGCAGACAGCTCAT 248
QY      81  LeuGlnIleAsnAsnLeuLeuSerGlnAspMetAlaThrTyrPheCysAlaArgTrpGlu 100
Db      249  TTGCAGATCAACAACCTCAAAAATGAGSACACGGCTACATATTTCTGTGGAAGCAATT 308
QY      101  ValTyrHisGlyTyrValProTyrTrpGlyGlnGlyThrThrValThrValSerSerGly 120
Db      309  ATTACCTACGCGGGTTTGCTAAGCTGGGGCCAGGAGACTGTGCTCATCTCTGCA--- 365
QY      121  GlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThrGln 140
Db      366  TCGAGCTCTCCGATCTTCATCTAGCGGTTCCAGCTCGACGATATCGTCATGACCCAG 425
QY      141  SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysAlaSer 160
Db      426  TCTCTTAATTCATGCTCCACCTCAGTGGAGACAGGGTCAAGCATCTCTGCAAGGCCAGT 485
QY      161  GlnAspValTyrAsnAlaValAlaTyrTyrGlnGlnLysProGlyGlnSerProLysLeu 180
Db      486  CAGAGTGTGAGTACGCTGTAGCTCGGTATCAACAAAACAGGAGCAATCTCCAAACTA 545
QY      181  LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200
Db      546  CTGATTTAATCTGACATCCACCGGACACCTGAGCTCCCTGATCCCTCACAGGAGTGA 605
QY      201  SerGlyProAspPheThrPheThrIleSerSerValGlnIleAsnLeuAlaValTyr 220
Db      606  TCTGGGACAGATTACTCTTCACCATACAGACGTGTGAGGCTTGAAGACCTGGCACTTCAT 665
QY      221  PheCysGlnGlnHisPheArgTyrProPheThrPheGlySerGlyTyrHisLeuGlnIle 240
Db      666  TACTGTGAGCAACATTATAGAGTCCGTAACGTTCCGAGGGGGAGCCAAAGCTGGAATA 725
QY      241  Lys 241
Db      726  AAA 728

RESULT 12
US-08-461-184-7
; Sequence 7, Application US/08461184
; Patent No. 5631158
; GENERAL INFORMATION:
; APPLICANT: DORAI, HAIMANTI
; APPLICANT: OPERMANN, HERMANN
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR HIGH PROTEIN
; TITLE OF INVENTION: PRODUCTION FROM NON-NATIVE DNA
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES, INC
; STREET: 45 SOUTH STREET
; CITY: HOPKINTON
; STATE: MA
; COUNTRY: USA
; ZIP: 07148
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,184
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/143,498

```

```

; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: KELLEY, ROBIN D
; REGISTRATION NUMBER: 34,637
; REFERENCE/DOCKET NUMBER: CRP093
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/248-7000
; TELEFAX: 617/248-7100
; INFORMATION FOR SRO ID NO.: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 909 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..752
; OTHER INFORMATION: /product= "741F6 sfv"
US-08-461-184-7

Alignment Scores:
Pred. No.: 1,84e-85 Length: 909
Score: 981.50 Matches: 185
Percent Similarity: 85.89% Conservative: 22
Best Local Similarity: 76.76% Mismatches: 33
Query Match: 41.86% Indels: 1
DB: 1 Gaps: 1

US-09-596-774-7 (1-443) x US-08-461-184-7 (1-909)
QY      1  GlnIleGlnLeuValGlnSerGlyProGluLeuLysPheProGlyGlnThrValLysIle 20
Db      9  GAATTCATTGGTGTGACGTCTGACCTGAGCTGAGAACCTGGAGAGACATGCAATC 68
QY      21  SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTrpValLysGlnAla 40
Db      69  TCTGCAAGGCTTCTGGGTATACCTTCAACAACTATGAAATGAATGAGGTGGAAGCGCT 128
QY      41  ProGlyGlnGlyLeuLysTrpMetGlyTyrPheAsnThrSerThnGlyGluSerThrPhe 60
Db      129  CCAGGAAGGCTTTAAAGTGGATGGGCTGATTAACACCACTGGAAGCCAAACATAT 188
QY      61  AlaAspAspPheLeuGlyValArgPheAspPheSerLeuGlnThrSerAlaAsnThrAlaTyr 80
Db      189  GCTGAAGAGTTCAAGGAGCGTTTGGCTTCTCTTGGAAACCTCTGCAGACAGCTCAT 248
QY      81  LeuGlnIleAsnAsnLeuLysSerGlnAspMetAlaThrTyrPheCysAlaArgTrpGlu 100
Db      249  TTGCAGATCAACAACCTCAAAAATGAGSACACGGCTACATATTTCTGTGGAAGCCAAATT 308
QY      101  ValTyrHisGlyTyrValProTyrTrpGlyGlnGlyThrThrValThrValSerSerGly 120
Db      309  ATTACCTACGCGGGTTTGCTAAGCTGGGGCCAGGAGACTGTGCTCATCTCTGCA--- 365
QY      121  GlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThrGln 140
Db      366  TCGAGCTCTCCGATCTTCATCTAGCGGTTCCAGCTCGACGATATCGTCATGACCCAG 425
QY      141  SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysAlaSer 160
Db      426  TCTCTTAATTCATGCTCCACCTCAGTGGAGACAGGGTCAAGCATCTCTGCAAGGCCAGT 485
QY      161  GlnAspValTyrAsnAlaValAlaTyrTyrGlnGlnLysProGlyGlnSerProLysLeu 180
Db      486  CAGAGTGTGAGTACGCTGTAGCTGGATCAACAAAACAGGAGCAATCTCTTAAACTA 545
QY      181  LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200
Db      546  CTGATTTAATCTGACATCCACCGGACACCTGAGTCCGATCGCTTCAAGGAGCTGA 605
QY      201  SerGlyProAspPheThrPheThrIleSerSerValGlnIleGlnAspLeuAlaValTyr 220

```

Db 606 TCTGGACAGATATTAATCTTCACCATCAGACGTGTCAGAGCTGAAGACCTGCACTTCAT 665
Qy 221 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyThrLysLeuGlnIle 240
Db 666 TACTGTGACGAACTATTAAGTGGCCGTACACGTTTCGAGAGGGGACCAAGCTGGAGATA 725
Qy 241 Lys 241
Db 726 AAA 728

RESULT 13

US-08-463-675-7

Sequence 7, Application US/08463675

Patent No. 5658763

GENERAL INFORMATION:

APPLICANT: DORAI, HAIMANTI

APPLICANT: OPPERMANN, HERMANN

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR HIGH PROTEIN

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESS: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES, INC

STREET: 45 SOUTH STREET

CITY: HOPKINTON

STATE: MA

COUNTRY: USA

ZIP: 07148

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/463,675

FILING DATE: 05-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/143,498

FILING DATE: 25-OCT-1993

ATTORNEY/AGENT INFORMATION:

NAME: KELLEY, ROBIN D

REGISTRATION NUMBER: 34,637

REFERENCE/DOCKET NUMBER: CRP093

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617/248-7000

TELEFAX: 617/248-7100

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 909 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: CDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 3..752

OTHER INFORMATION: /product= "741F8 sfv."

US-08-463-675-7

Alignment Scores: 1.84e-85 Length: 909
Pred. No.: 981.50 Matches: 185
Score: 85.89% Conservative: 22
Percent Similarity: 76.76% Mismatches: 33
Best Local Similarity: 41.86% Indels: 1
Query Match: 1 Gaps: 1

US-09-596-774-7 (1-443) x US-08-463-675-7 (1-909)

Qy 1 GlnIleGlnLeuValGlnSerGlyProGlnLeuLysLysProGlyGlnThrValLysIle 20
Db 9 GAGATCCAAATGTGGAGCTGAGACCTGAGCTGAAGAGCCTGGAGAGACAGTCAAGATC 68

Qy 21 SerCysLysAlaSerGlyThrProPheThrAsnTyrGlyMetAsnTyrValLysGlnAla 40
Db 69 TCTGCAAGAGCTTCTGGGTATTAATCTTCACCAACTATGAAATGAATGAGTGGTGAACAGAGCT 128
Qy 41 ProGlyGlnGlnLysLeuLysTyrPheMetGlyTyrPheAsnThrSerThrGlyLysSerThrPhe 60
Db 129 CCAGAAAGGGGTTTAAAGTGGATGGCTGGTAAACCAACACTGAGAGACCAACATAT 188
Qy 61 AlaAspAspPheLysGlyArgPheAspPheSerLeuGlnThrSerAlaAsnThrAlaTyr 80
Db 189 GCTGAAGAGTTCAGAGGACGGATTGCTTCTTTGAAACCTTGGCAGACAGTGGCTAT 248
Qy 81 LeuGlnIleAsnAsnLeuLysSerGluAspMetAlaThrTyrPheCysAlaArgTyrGln 100
Db 249 TTGCAATCAACAACTCCAAATAATGAGACACGGCTCATATTTCTGTGAAGCAATTT 308
Qy 101 ValTyrHisGlyTyrValProTyrTyrPheGlnGlnLysThrValThrValSerSerGly 120
Db 309 ATTACTACGGCGGGTTTGGTAACTGGGGCCAGGAGCTGTGCTGCTGCTGCA--- 365
Qy 121 GlyGlyLysSerGlyGlyGlySerGlyGlyGlyLysSerAspIleGlnLeuThrGln 140
Db 366 TCGAGCTCTCCGATCTTCACTTAGCGGTTCCAGCTCGAGCATGATGTCATACCAAG 425
Qy 141 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSer 160
Db 426 TCTCTTAATTCATGCTCACTGATGAGGAGACAGGCTCAGCATCTCTGCAAGGCCAGT 485
Qy 161 GlnAspValTyrAsnAlaValAlaTyrTyrGlnGlnLysProGlyGlnSerProLysLeu 180
Db 486 CAGATGTGAGTACTGCTGTAAGCTGATCAACAAACACAGGAGCAATCTCTAACTA 545
Qy 181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200
Db 546 CTGATTTACTGACATCCACCGGACACCTGAGTCTCTGATCTCTTACACAGGAGTGA 605
Qy 201 SerGlyProAspPheThrPheThrIleSerSerValGlnIleGlnAspLeuAlaValTyr 220
Db 606 TCTGGACAGATTAATCTTCACCATCAGAGTGTGAGGCTGAAGACCTGGCACTTCAT 665
Qy 221 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyThrLysLeuGlnIle 240
Db 666 TACTGTGACGAACTATTAAGTGGCCGTACACGTTTCGAGAGGGGACCAAGCTGGAGATA 725
Qy 241 Lys 241
Db 726 AAA 728

RESULT 14

US-08-464-589-7

Sequence 7, Application US/08464589

Patent No. 5733782

GENERAL INFORMATION:

APPLICANT: DORAI, HAIMANTI

APPLICANT: OPPERMANN, HERMANN

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR HIGH PROTEIN

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES, INC

STREET: 45 SOUTH STREET

CITY: HOPKINTON

STATE: MA

COUNTRY: USA

ZIP: 07148

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/464,589

FILING DATE: 05-JUN-1995

```
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 08/143,498
/ FILING DATE: 25-OCT-1993
/ ATTORNEY/AGENT INFORMATION:
/ NAME: KELLEY, ROBIN D
/ REGISTRATION NUMBER: 34,637
/ REFERENCE/DOCKET NUMBER: CRP093
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617/248-7000
/ TELEFAX: 617/248-7100
/ INFORMATION FOR SEQ ID NO: 7:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 909 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 3..752
/ OTHER INFORMATION: /product= "741F8 sFV"
US-08-464-589-7

Alignment Scores:
Pred. No.: 1,84e-85 Length: 909
Score: 981.50 Matches: 185
Percent Similarity: 85.89% Conservative: 22
Best Local Similarity: 76.76% Mismatches: 33
Query Match: 41.86% Indels: 1
DB: 1 Gaps: 1

US-09-596-774-7 (1-443) x US-08-464-589-7 (1-909)

Qy 1 GlnlGlnLeuValGlnSerGlyProGluLeuLysProGlyGluThrValLysIle 20
Db 9 GAGATCCAAATGGGAGCTGACCTGAGCTGAGAAAGCCTGAGAGACGTCMAAGTC 68
Qy 21 SerCysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTPValLysGlnAla 40
Db 69 TCCGCAAGGCTTGGGTATACCTTCACAACTATGGAATGAACTGGGTGAAACAGGCT 128
Qy 41 ProGlyGlnGlyLeuLysTyrPheGlyTyrPheLeuAsnThrSerThrGlyLysThrPhe 60
Db 129 CCAGGAAAGGCTTAAAGTGGATGGCTGATTAACCAACAACCTGAGAGCCAACTAT 188
Qy 61 AlaAspAspPheLysGlyTyrGlyPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80
Db 189 GCTGAAGAGTTCAAGGAGCAGGCTTGCCTTCCTTGGAAACCTCTGCCAGCAGCTGCTAT 248
Qy 81 LeuGlnIleAsnAsnLeuLysSerGlyAspPheAlaThrTyrPheCysAlaArgTyrGlu 100
Db 249 TTGCAGATCAACAACCTCAAAAATGAGAGACAGGCTACATATTCGTGGAAGGCAATT 308
Qy 101 ValTyrHisGlyTyrValProTyrTyrPylGlnGlyThrThrValThrValSerSerGly 120
Db 309 ATTACCTACGCGCGGCTTGGCAACTGGGGCCAGAGCACTGCGTCACTGCTGCA--- 365
Qy 121 GlyLysLysSerGlyGlyGlyLysSerGlyGlyLysSerAspIleGlnLeuThrGln 140
Db 366 TCGAGCTCCTCCGATCTTCATCTAGCGGTTGCCAGCTCGAGCATATCGCATGACCCAG 425
Qy 141 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSer 160
Db 426 TCTCTAAATTCATGCTCAGCTCAGTGGAGAGAGGGGTCCAGCATCTCTCCAAAGGCCAGT 485
Qy 161 GlnAspValTyrAsnAlaValAlaTyrTyrGlnGlnLysProGlyLysSerProLysLeu 180
Db 486 CAGAGTGTGAGTACTGCTGAGCTGTATCAACAAACCAAGGCGCAATCTCTAAACTA 545
Qy 181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200
Db 546 CTGATTTTACGACATCCACCCGCGACACGAGTCCCTGATGCTTCACAGGACAGTGA 605
```

```
Qy 201 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 220
Db 606 TCTGGACAGATTATCTCATCCATCAGCAGTGTGAGGCTGAAGACCTGGCACTTCAT 665
Qy 221 PheCysGlnGlnHisPheArgTyrProPheThrPheGlySerGlyThrLysLeuGluIle 240
Db 666 TACTGTAGCAGCACTTATTAAGTGGCCGTACACGTTCCGAGGGGGACCAAGCTGGAGATA 725
Qy 241 Lys 241
Db 726 AAA 728

RESULT 15
US-08-461-838-1
/ Sequence 1, Application US/08461838
/ Patent No. 5753204
/ GENERAL INFORMATION:
/ APPLICANT: Huston, James S.
/ APPLICANT: Oppermann, Hermann
/ APPLICANT: Houston, L. L.
/ TITLE OF INVENTION: Ring, David B.
/ TITLE OF INVENTION: Biosynthetic Binding Proteins For
/ NUMBER OF SEQUENCES: 11
/ CORRESPONDENCE ADDRESS:
/ ADDRESSER: Testa, Hurwitz & Thideault/Patent Department
/ STREET: Exchange Place, 53 State Street
/ CITY: Boston
/ STATE: Massachusetts
/ COUNTRY: USA
/ ZIP: 02109
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent In Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/461,838
/ FILING DATE:
/ CLASSIFICATION: 424
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Kelley, Robin D.
/ REGISTRATION NUMBER: 34,637
/ REFERENCE/DOCKET NUMBER: 2054/22
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 617-248-7477
/ TELEFAX: 617-248-7100
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 909 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 3..752
/ OTHER INFORMATION: /product= "741F8 sFV" C-terminal
/ OTHER INFORMATION: Gly4-Cys"
US-08-461-838-1

Alignment Scores:
Pred. No.: 1,84e-85 Length: 909
Score: 981.50 Matches: 185
Percent Similarity: 85.89% Conservative: 22
Best Local Similarity: 76.76% Mismatches: 33
Query Match: 41.86% Indels: 1
DB: 1 Gaps: 1

US-09-596-774-7 (1-443) x US-08-461-838-1 (1-909)

Qy 1 GlnlGlnLeuValGlnSerGlyProGluLeuLysProGlyGluThrValLysIle 20
```

```
Db      ::::|||||
9      GAGATCCCAATTGGCTGCACTGCACTGCAAGAAAGCCTGAGAGACAGTCAAGATC 68
QY      21  SerCysIysAlaSerGlyThrProPheThrAntYrGlyMetAsnTrpValIysGlnAla 40
Db      69  TCCTGCAAGGCTTCGGGTATACCTTCACAACTATGAAATGAACTGGGTGAAGCAGCT 128
QY      41  ProGlyGlnGlyLeuIleYrTrpMetGlyTrpIleAsnThrSerThrGlyIleSerThrPhe 60
Db      129  CCAGGAAAGGTTTAAAGTGAATGGGTGGATTAACCAACCAACACTGAGAGCCACATAT 188
QY      61  AlaAspAspPheIysGlyArgPheAspPheSerIleuGluThrSerAlaAsnThrAlaTrp 80
Db      189  GCTGAAGAGTTCAAGGAGCGGTTTGCCTTCCTTTGAAACCTTGCCAGCAGCTGCTAT 248
QY      81  LeuGlnIleAsnIleuIleSerGlyLeuAspMetAlaThrTrpPheCysAlaArgTrpGlu 100
Db      249  TTGCAGATCAACMACCTCAAAATGAGACACGCTACATATTTCTGTGAAGGCAATT 308
QY      101  ValTyriAsgIyTyValProTyTrpGlyGlnIyThrValThrValSerSerGly 120
Db      309  ATTACCTACGGCGGTTTGCTACTGAGGCCAAGGAACTTGCTCATCTGTCTTCA--- 365
QY      121  GlyIyGlySerGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThrGln 140
Db      366  TCGAGCTCCTCGGATCTTCATCTAGCGGTTCCAGCTCAGCGATATCGTATGACCCAG 425
QY      141  SerIleYrPheLeuSerThrSerValGlyAspArgValSerIleThrCysIysAlaSer 160
Db      426  TCTCCTTAATTCATGTCCACGTCAGTGGAGACAGGGTCAGCATCTCTGCAAGGCCAGT 485
QY      161  GlnAspValTyriAsnAlaValAlaTrpIyGlnGlnIyProGlyIyIleSerProIySer 180
Db      486  CAGATGTGAATACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 545
QY      181  LeuIleYrSerAlaSerSerArgTrpThrGlyValProSerArgPheThrGlySerGly 200
Db      546  CTGATTTACTGACATCCACCGGACACTGGAGTCCCTGATCGCTTCACAGGCAGTGSA 605
QY      201  SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTy 220
Db      606  TCTGGGACAGATTTACTCTCACCATCAGCAGTGTGACAGGCTGAGACCTGGCATTGAT 665
QY      221  PheCysGlnGlnIleAspPheArgThrProPheThrPheGlySerGlyThrIySerGluIle 240
Db      666  TACTGTCAACAATTATAGAGTCCGTACACGTTTGGAGGGGGGACCAAGCTGAGATA 725
QY      241  Lys 241
Db      726  AAA 728
```

Search completed: April 20, 2005, 16:41:44
Job time : 239.44 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 20, 2005, 10:19:55 ; Search time 758.098 Seconds

(without alignments)
3545.124 Million cell updates/sec

Title: US-09-596-774-7

Perfect score: 2345

Sequence: 1 QIOLVQSGPELKKPGEIVKI.....LSTATKDYDALHMQTLAPR 443

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 5622541 seqs, 303355566 residues

Total number of hits satisfying chosen parameters: 11245082

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=x1h
-Q=/cg2_1/USPTO.spool/h/US09596774/runat_20042005_075444_12560/app.query.fasta_1.1230
-DB=Published Applications NA -QPM=fastcap -SUFFIX=apric20.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blom62
-TRANS=human40.cdd -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09596774@CGN_1.698@runat_20042005_075444_12560
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSELOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:*

1: /cg2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cg2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cg2_6/ptodata/1/pubpna/US06_NEW_PUB.seq:*
4: /cg2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cg2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cg2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cg2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cg2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cg2_6/ptodata/1/pubpna/US09_PUBCOMB.seq:*
10: /cg2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cg2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cg2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cg2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cg2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cg2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cg2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cg2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
18: /cg2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq:*
19: /cg2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
20: /cg2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
21: /cg2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
22: /cg2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. NO. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1249	53.3	1920	US-10-120-198B-1	Sequence 1, Appl1
2	1249	53.3	6834	US-10-120-198B-5	Sequence 5, Appl1
3	1015.5	43.3	7654	US-10-006-773-1	Sequence 1, Appl1
4	1015.5	43.3	7654	US-10-006-773A-1	Sequence 1, Appl1
5	994	42.4	771	US-10-013-173-38	Sequence 38, Appl1
6	994	42.4	771	US-10-150-762-38	Sequence 38, Appl1
7	994	42.4	771	US-10-244-821-38	Sequence 38, Appl1
8	981.5	41.9	909	US-09-887-853-1	Sequence 1, Appl1
9	969.5	41.3	909	US-09-887-854-1	Sequence 1, Appl1
10	930	39.7	1350	US-08-812-381-1	Sequence 1, Appl1
11	928	39.6	1350	US-09-774-681-1	Sequence 1, Appl1
12	917.5	39.1	1515	US-10-239-656-78	Sequence 78, Appl1
13	913	38.9	816	US-09-766-543-9	Sequence 9, Appl1
14	903.5	38.5	723	US-10-127-890-90	Sequence 90, Appl1
15	903.5	38.5	723	US-10-717-243-90	Sequence 90, Appl1
16	897	38.3	840	US-09-766-543-11	Sequence 11, Appl1
17	889.5	37.9	1626	US-10-071-485-84	Sequence 84, Appl1
18	888.5	37.9	804	US-10-071-485-84	Sequence 84, Appl1
19	888.5	37.9	2133	US-10-071-485-89	Sequence 89, Appl1
20	886.5	37.8	1509	US-10-239-656-74	Sequence 74, Appl1
21	875.5	37.3	10511	US-10-059-261-109	Sequence 109, Appl1
22	875.5	37.3	10511	US-10-627-649-109	Sequence 109, Appl1
23	872.5	37.2	729	US-10-610-452-9	Sequence 9, Appl1
24	872.5	37.2	729	US-10-610-452-9	Sequence 9, Appl1
25	860.5	36.7	761	US-10-354-246-4	Sequence 4, Appl1
26	859	36.6	1869	US-10-378-832A-1	Sequence 1, Appl1
27	854.5	36.4	1509	US-10-239-656-76	Sequence 76, Appl1
28	854	36.4	777	US-10-096-246-5	Sequence 5, Appl1
29	854	36.4	777	US-10-096-246-7	Sequence 7, Appl1
30	849.5	36.2	1497	US-10-239-656-72	Sequence 72, Appl1
31	848	36.2	702	US-10-071-485-92	Sequence 92, Appl1
32	846	36.1	777	US-10-096-246-6	Sequence 6, Appl1
33	841.5	35.9	756	US-09-984-186-17	Sequence 17, Appl1
34	841.5	35.9	756	US-10-237-667-17	Sequence 17, Appl1
35	841.5	35.9	756	US-10-237-708-17	Sequence 17, Appl1
36	841.5	35.9	756	US-10-237-866-17	Sequence 17, Appl1
37	841.5	35.9	756	US-10-237-871-17	Sequence 17, Appl1
38	841.5	35.9	756	US-10-237-624-17	Sequence 17, Appl1
39	841.5	35.9	756	US-10-702-536-17	Sequence 17, Appl1
40	841.5	35.9	756	US-10-702-636-17	Sequence 17, Appl1
41	830.5	35.4	739	US-09-887-853-5	Sequence 5, Appl1
42	830.5	35.4	739	US-10-683-547-5	Sequence 5, Appl1
43	830.5	35.4	774	US-10-096-246-3	Sequence 3, Appl1
44	827.5	35.3	723	US-10-335-290-169	Sequence 169, Appl1
45	826	35.2	921	US-10-363-349-12	Sequence 12, Appl1

ALIGNMENTS

RESULT 1
US-10-120-198B-1
Sequence 1, Application US/10120198B
Publication No. US20030215427A1
GENERAL INFORMATION: Jensen, Michael
APPLICANT: Jensen, Michael
TITLE OF INVENTION: CE7-SPECIFIC REDIRECTED IMMUNE CELLS
FILE REFERENCE: 1954-337
CURRENT APPLICATION NUMBER: US/10/120, 198B
CURRENT FILING DATE: 2002-04-11
PRIOR APPLICATION NUMBER: 60/282, 859
PRIOR FILING DATE: 2001-04-11
NUMBER OF SEQ ID NOS: 11
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1
LENGTH: 1920
TYPE: DNA
ORGANISM: artificial sequence
FEATURE:
OTHER INFORMATION: mouse-human chimera

FEATURE:
 NAME/KEY: CDS
 LOCATION: (14) (1906)
 OTHER INFORMATION: seqvfc construct
 US-10-120-198B-1

Alignment Scores:
 Pred. No.: 1.19e-120 Length: 1920
 Score: 1249.00 Matches: 276
 Percent Similarity: 54.60% Conservative: 62
 Best Local Similarity: 44.59% Mismatches: 96
 Query Match: 53.26% Indels: 186
 Gaps: 12

US-09-596-774-7 (1-443) x US-10-120-198B-1 (1-1920)

```

QY 1 GlnIleGlnLeuValGlnSerGlyProGluLeuLysProGlyLysThrValLysIle 20
Db CAGGTCCAACTGCAGCAGCCTGGGGCTGACTGGTGAAGCCTGGGGCTTCAGTGAAGCTG 139
QY 21 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetLeuThrValLysGlnAla 40
Db TCCTGCAAGGCTTCTGCTGACACCTTCAACGGCTACCTGAGATCACTGGTGAACCAAGG 199
QY 41 ProGlyGlnGlyLeuLysTyrPheGlyTyrPheLeuThrSerThrGlyLysThrPhe 60
Db CCTGACATGGCTTGAAGTGAAGATTAATCTTACCAACGGCTGACTACTAC 259
QY 61 AlaSerPhePheLysGlyLysPheAspPheSerLeuGlnThrSerAlaThrAlaTyr 80
Db AATAAGAGTTCAAGAGCAAGCAACCTGACTGATGACAAATCCTCCACACAGCCTTC 319
QY 81 LeuGlnIleAsnLeuLeuLysSerGlyLysPheAlaThrTyrPheCysAlaLysTyrGlu 100
Db ATGACACTCAGCGGCTTCACTGAGACCTGACGCTTCTTCTGCAAGA----- 373
QY 101 ValTyrHisGly-----TyrValProTyrTyrGlyGlnGlyThrThrValThrVal 117
Db GATTACTACGGTACTAGCTACACTTGTGACTACTGGGGCCAAAGCACACTCTCACAGTC 433
QY 118 SerSerGlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlySerAspIleGln 137
Db TCCCTCAGAGGTGGCGTGAAGTGAAGTGGCGGATCCGGTGGCGGAGTGAAGCATCAG 493
QY 138 LeuThrGlnSerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCys 157
Db ATGACACAACTTTCATCTCTCTTCTGATCTTGAAGACAGAGTCAACATTACTTGC 553
QY 158 LysAlaSerGlnAspValTyrAsnAlaValAlaTyrTyrGlnGlnLysProGlyGlnSer 177
Db AAGGCTAATGAAGCATTAATAATCGGTAGCTGTATCAGACAGACACGAAATTTCT 613
QY 178 ProLysLeuLeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThr 197
Db CTTAGGCTCTTAATATCTGTGTCACCAATTGGTAACTGGGGTTCCTTCAAGATTGAGT 673
QY 198 GlySerGlySerGlyProAspPheThrPheThrIleSerSerValGlnIleGluAspLeu 217
Db GCGAGTGAAGTGAAGAGTATACCTCACTTCAACAGTCTTCAAGGCTGAAGATTTT 733
QY 218 AlaValTyrPheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyThrLys 237
Db GCTACTTATTAATCTGTCAACATATTTGAATCACTTCACTTCACTGCGCTCGGGAGAGAG 793
QY 238 LeuGluIleLys----- 241
Db CTGAGATCAAAAGTAGAACCCAAATCTTGTGACAAAATGACACATGCCACCGTGCCA 853
QY 241 ----- 241
Db GCACTGAACTCCCTGGGGGAGCGTCAAGTCTTCTCTTCCCCCAAAACCAAGACACC 913
QY 242 -----AlaLeuGluIleSer----- 246
  
```

```

Db 914 CTATGATCTCCGGAGCCCTGAGGTCAATGCGTGTGTGAGCGTGAACCAAGAAC 973
QY 246 ----- 246
Db 974 CTTGAGGTCAAGTTCAACTGTGATCGTGAACGGCGTGAAGTGAATATGCAAGACAAAG 1033
QY 247 -----AsnSerValMetTyrPheSerSerValProValIleGln 260
Db 1034 CCGCGGAGAGAGATCAACAGACGCTACCGTGTGCTCACTTCACTGCTGAC 1093
QY 261 -----LysValAsnSerThrThrThrLysPr 269
Db 1094 CAGGACTGCTGAATGCGAAGAGTACACTGCAAGGTCTCCAAACAAAGCCTCCACGCC 1153
QY 269 oValLeuArgThrProSerProValHisProThrGlyThrSerGlnProGlnArgProGln 289
Db 1154 CCCATCGAAGAAACCATCTCCA---AAGCCAAGGCGAGCCCGAAGAACACAGTGTAC 1210
QY 289 uAspCys----- 291
Db 1211 ACCCTGCAACCATCAGAGATGAGCTGACCAAGAACCAAGTCAAGCTGCTGCTGCTC 1270
QY 292 -----ArgProArgGlySerValLysGln 299
Db 1271 AAAGGCTTATCCCAAGCAGCATCGCGCTGAGTGGAGAGCAATGGCGACCGAGAAC 1330
QY 299 YThrGlyLeuAspPheLeuGlnAspProLysLeu---CysTyrLeuLeuAspGlyIleLe 318
Db 1331 AACT-----ACAAGACCAAGCGCTCCGCTGTGATCACC-GACGGCTCCTT 1374
QY 318 uPheIleTyr----- 321
Db 1375 CTTCCTCTACAGCAAGCTACCGTGAACAAGACAGTGGACAGAGGAGAAAGCTTCTTC 1434
QY 321 ----- 321
Db 1435 ATGCTCGATGATGATGAGGCTGTGCACAACACTACAGCAAGAAAGCTTCCTGTC 1494
QY 322 -----GlyValIleIle----- 325
Db 1495 TCCCGGAAATATGACCTGATATGTGCTGGGGCGCTGCGCGCTCTCTTTCATTGG 1554
QY 326 -ThrAlaLeuTyrLeuArgAlaLysPheSerArgSerAlaGluThrAlaAlaAsnLeuGln 345
Db 1555 GCTAGGACTCTTCTTCAAGTGAAGTTCAGACAGACGAGACGCCGCCGCTACACAGA 1614
QY 345 nAspProAsnGlnLeuTyrAsnGlnLeuAsnLeuGlnLysArgGluGluTyrAspValLe 365
Db 1615 GGGCAGAAACCAAGCTCTATTAACAGCTCAATCTAGGACGAAGAGAGATGATGTTT 1674
QY 365 uGluLysLysAlaArgAspProGluMetGlyLysGlnGlnArgTyrArgAsnPr 385
Db 1675 GGCACAAAGACGTTGGCGGAGCCCTTGAATGGGGGAAAG---CCGAGAAGAAAGAACCC 1731
QY 385 oGlnGlnGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaGluAlaTyrSerGluIle 405
Db 1732 TCGAGAGGCTGTATCAATGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1791
QY 405 eGlyThrLysGlyLysArgArgArgGlyLysGlyHisAspGlyLeuTyrGlnGlyLeuSe 425
Db 1792 TGGAGTGAAGGAGCGCGCGGAGGGGCAAGGGGCAAGTGGCTTTCACAGGCTCGAG 1851
QY 425 rThrAlaThrLysAspThrTyrAspAlaLeuHisMetGlnThrLeuAlaProArg 443
Db 1852 TACAGCCAAAGAGACCTTACAGCGCTTCAATGACAGGCCCTGCCCCCTCGC 1906
  
```

RESULT 2

US-10-120-198B-5
 ; Sequence 5, Application US/10120198B
 ; Publication No. US20030215427A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Jensen, Michael

; TITLE OF INVENTION: CEF7-SPECIFIC REDIRECTED IMMUNE CELLS
 ; FILE REFERENCE: 1954-337
 ; CURRENT APPLICATION NUMBER: US/10/120,198B
 ; CURRENT FILING DATE: 2002-04-11
 ; PRIOR APPLICATION NUMBER: 60/282,859
 ; PRIOR FILING DATE: 2001-04-11
 ; NUMBER OF SEQ ID NOS: 11
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 5
 ; LENGTH: 6834
 ; TYPE: DNA
 ; ORGANISM: artificial sequence
 ; FEATURE:
 ; OTHER INFORMATION: plasmid construct
 US-10-120-198B-5

Alignment Scores:

Prod. No.:	6,65e-120	Length:	6834
Score:	1249.00	Matches:	276
Percent Similarity:	54.60%	Conservative:	62
Best Local Similarity:	44.59%	Mismatches:	96
Query Match:	53.26%	Indels:	186
DB:	17	Gaps:	12

US-09-596-774-7 (1-443) x US-10-120-198B-5 (1-6834)

QY 1 GlnIleGlnLeuValGlnSerGlyProGluLeuLysProGlyGluThrValLysIle 20
 DB 1878 CAGGTCACACGACGACGCTGGGCTGAAGCTGGAAGCTGGGGCTTCAGTGAAGCTG 1937
 QY 21 SerCysValAspAspGlyTyrProPheThrAsnTyrGlyMetAsnTyrValLysGlnAla 40
 DB 1938 TCCTGCAAGGCTTCTGGCTACCTTCAACGGCTACGATGCACTGGGTGAAGCAGAGG 1997
 QY 41 ProGlyGlnGlyLeuLeuTyrPheGlyTyrIleAsnThrSerThrGlyGluSerThrPhe 60
 DB 1998 CCTGACACATGGCTTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 2057
 QY 61 AlaAspAspPheLysGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80
 DB 2058 AATGAGAGCTTCAGAGCAGAGCAGCAGCTGAGTGAAGTGAAGTGAAGTGAAGTGAAG 2117
 QY 81 LeuGlnIleAsnAsnLeuLysSerGluAspMetAlaThrTyrPheCysAlaArgTyrPhe 100
 DB 2118 ATGCACTCAGCGGCTGACATCTGAGACCTGTCGACGCTAATTTCTGTCAGAGA----- 2171
 QY 101 ValTyrHisGly-----TyrValProTyrTyrGlyGlnGlyTyrThrValThrVal 117
 DB 2172 GATTACTGCGTACTAGCTACCAACTTGTACTGAGGCGCAAGGCGACCACTCTCAGAGT 2231
 QY 118 SerSerGlyGlyGlySerGlyGlyGlySerGlyGlyGlySerGlyGlySerAspIleGln 137
 DB 2232 TCCTCAGAGGCTGGGCTGAGTGAAGTGGCGGATCCGGTGGCGGAGTGAAGTGAAGTGAAG 2291
 QY 138 LeuThrGlnSerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCys 157
 DB 2292 ATGACACATCTTCACTCTCTCTTCTGTATCTCTGAGAGACAGAGTCAACCTTACTTGC 2351
 QY 158 LysAlaSerGlnAspValTyrAsnAlaValAlaTyrTyrGlnGlnLysProGlyGlnSer 177
 DB 2352 AAGCTAATGAAACATTAATTAATCGGTATGCTGATCAGACAGACACCAAGAAATCT 2411
 QY 178 ProLysLeuLeuLysSerLysSerArgTyrThrGlyValProSerThrPheThr 197
 DB 2412 CCTAGGCTCTTAATATCTGCTGCAACCAATTGTAACCTGGGCTCTTCAAGATTCAGT 2471
 QY 198 GlySerGlySerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeu 217
 DB 2472 GCGAGTGAATCTGGAAGATTAACCTCAGCATTACAGCTTCAAGCTGAAGATTTT 2531
 QY 218 AlaValTyrPheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyTyrLys 237
 DB 2532 GCTACTTATTAATGCAACAAATATTGAGTATTCATTCAGCTTCAAGCTGAGGAGCAGAG 2591

QY 238 LeuGlnIleLys----- 241
 DB 2592 CTCGAGATCAAGATGAAGAACCCCAATCTTCTGACAAAACACTCACATGCCACCGGCCCA 2651
 QY 241 ----- 241
 DB 2652 GCACCTGAATCTGGGGGACCGTCAAGCTTCTCTTCCCCCAAAACCAAGAGACACC 2711
 QY 242 -----AlaLeuGlnLysSer----- 246
 DB 2712 CTCATGATCTCCGGACCCCTGAGGTCAATGCTGTGTGAGACGTGAGCCAGAGAGAC 2771
 QY 246 ----- 246
 DB 2772 CCTGAGCTCAAGTTCACTGCTAGCTGACCGCGTGAAGTGCATATGCCAAGACAAAG 2831
 QY 247 -----AsnSerValMetTyrPheSerSerValProValLeuGln 260
 DB 2832 CCGGGAGAGAGAGTACACAGACAGTACCGTGTGCTGCTCCTCAGCGTCTGAC 2891
 QY 261 -----LysValAsnSerThrThrThrLysPhe 269
 DB 2892 CAGGACTGGCTGAATGGCAGAGTACAGAGTCAAGGCTCTCCAAACCAAGCCCTCCAGCC 2951
 QY 269 ovalLeuArgThrProSerProValHisProThrGlyTyrSerGlnProGlnArgProGln 289
 DB 2952 CCATTCAGAAACCATCTTCCA--AACCAAGGGCAGCCCCCAGAAACACAGGTGAC 3008
 QY 289 uAspCys----- 291
 DB 3009 ACCCTGCACATCAGACAGATGAGCTGACCAAGAACAGTCAAGCTGACCTGCTGTC 3068
 QY 292 -----ArgProArgGlySerValLysGly 299
 DB 3069 AAAGCTTCTATCCAGAGCATGCGCTGAGTGGAGAGCAATGGACACCGAGAAC 3128
 QY 299 YThrGlyLeuAspPheLeuGluAspProLysLeu-----CysTyrLeuLeuAspGlyIleLe 318
 DB 3129 AAC-----ACAAAGCAGCGCTCCGCTGCTGAGCTCC-GACGGCTCTT 3172
 QY 318 uPheIleTyr----- 321
 DB 3173 CTTCCTCTACAGCAAGCTCACCGTGAACAAGACAGAGTGGCAGAGGAGAACTTCTTC 3232
 QY 321 ----- 321
 DB 3233 ATGCTCCGTATGATGAGGCTCTGCAACCACTACACGAGAAAGCTCTCTGTC 3292
 QY 322 -----GlyValIleIle----- 325
 DB 3293 TCCGGGAGAAATGGCCCTGATTTGCTGGGGGGCTGCCCGGCTCTCTTTTCATTGG 3352
 QY 326 -ThrAlaLeuTyrLeuArgAlaLysPheSerArgSerAlaGluThrAlaAlaAsnLeuGln 345
 DB 3353 GCTAGGATCTTCTTCAAGATGAAGTTCAAGAGAGGAGGAGACCCCGCGGTACAGCA 3412
 QY 345 nAspProAsnGlnLeuTyrAsnGlnLeuAsnLeuGlyArgArgGlnGluTyrAspValLe 365
 DB 3413 GGGCCAGAACCGCTCTAATTAACGAGCTCAATCTGAGCAGAAAGAGAGTGCAGTGT 3472
 QY 365 uGluLysLysArgAlaArgAspProGluMetGlyGlyLysGlnGlnArgArgAsnPro 385
 DB 3473 GGAACAGAGAGTGGCGGAGCCCTGAGATGGGGGAGAAAG---CCGAGAGAGAGAACCC 3529
 QY 385 oGlnGlnGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaGlnAlaTyrSerGlnI 405
 DB 3530 TCAGAAAGGCTGTACATGAATGACAGAAAGATGAATGCGGAGGCTTACAGTGAAT 3589
 QY 405 eGlyThrLysGlyGluArgArgArgGlyLysGlyHisAspGlyLeuTyrGlnGlyLeuSe 425
 DB 3590 TGGATGAAGAGGCGAGGCGCGAGGGGAGGAGGAGAGAGTGGCTTTTACAGAGGCTTCAG 3649

```

Qy 425 rThralaThrlYsAspThrlYrAspAlaLeuHlSwetGlnThrlLeuAlaProArg 443
Db 3650 TACAGCCACCAAGAGACCACTACGAGCGCCCTTCACTGACGCGCCCTCCGC 3704

RESULT 3
US-10-006-773-1
; Sequence 1, Application US/10006773
; Publication No. US20020132983A1
; GENERAL INFORMATION:
; APPLICANT: Junghans, Richard P.
; TITLE OF INVENTION: Antibodies as Chimeric Effector Cell Receptors Against Tumor Anti-
; FILE REFERENCE: 003
; CURRENT APPLICATION NUMBER: US/10/006,773
; CURRENT FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: 60/250,089
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 7654
; TYPE: DNA
; ORGANISM: Homo sapiens and Mus sp.
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (2428)..(3759)
; OTHER INFORMATION: Chimeric IgTCR sequence contained in retroviral vector. Retrovi-
; OTHER INFORMATION: al vector sequence (non-coding regions) are incidental to the inv-
; OTHER INFORMATION: ention. The translated (coding region) is relevant to the invent-
; OTHER INFORMATION: ion. (pertinent to Figure 3.)
US-10-006-773-1

Alignment Scores:
Pred. No.: 2.53e-95 Length: 7654
Score: 1015.50 Matches: 225
Percent Similarity: 57.74% Conservative: 25
Best Local Similarity: 51.96% Mismatches: 58
Query Match: 43.30% Indels: 125
DB: 13 Gaps: 8

US-09-596-774-7 (1-443) x US-10-006-773-1 (1-7654)
Qy 131 GLyglYglYSerApplIegInLeuThrGlnSerHlSlyPheLeuSerThrSerValGly 150
Db 2473 GGTTCACCTCCAGCATCCAGCTGACCCGACCCAGCCAGCCAGCCGCGCGCTGGGT 2532
Qy 151 AspArgValSerIleThrCysAlaSerGlnAspValTyrAsnAlaValAlaTProTyr 170
Db 2533 GACGAGTGAACATCACTGTAAGCCAGTCAAGATGTGGTACTTCTGTAGCTTGATAC 2592
Qy 171 GlnGlnLysProGlyGlnSerProLysLeuLeuIleTyrSerAlaSerSerArgTyrThr 190
Db 2593 CAGCAGAGGCGAGGTAAAGCTCCAAAGCTGCTGATCTAATCGACATCAACCCGCAACT 2652
Qy 191 GLyAlaProSerArgRhetHrGlnSerGlySerGlyProAspPheThrPheThrIleSer 210
Db 2653 GGTGTGCAAGCAATTCAGGGTGAAGGCGGTAGCGGTACCACTTCACTTCCATCAGCATGAC 2712
Qy 211 SerValGlnAlaGlnAspLeuAlaValTyrPheCysGlnGlnHis-----PheArgThr 228
Db 2713 AGCCTCCAGCCAGAGGACATCGCACTACTACGCGCAAAATTAAGCTCTATCGG--- 2769
Qy 229 ProPheThrPheGlySerGlyThrlYrLeuGlnIleLys----- 241
Db 2770 -----TCGTTCCGCAAGGAGCAAGAGTGAATCAAAAGAGGTGCTCAGATCGCGT 2823
Qy 241 ----- 241
Db 2824 GATCCGAGCTCTGTGGCTCAGATCGAGTCCAACTGTGTGAGAGCGGTGAGAGTGT 2883
Qy 242 -----AlaLeuGlnIleSerAsnSerVal----- 249
Db 2884 GTGCAACTGCGCGGTCTGCGCTGTCTGCTCCGATCTGAGCTTTCATTTCCACCACA 2943

```

```

Qy 249 ----- 249
Db 2944 TATTGATGAGTTGGTGAGACAGGCACTGGAAAGGCTTGAGTGGATTGAGAAATT 3003
Qy 249 ----- 249
Db 3004 CATCCAGATAGACATGATTAATCTATGCGCGCTCTTAAGAGTATGATTACATATCG 3063
Qy 249 ----- 249
Db 3064 CGAGCAACGCCAAGAACACATGTTCTTCGAATAGACAGACCTTGAGACCCGAGACACC 3123
Qy 250 -----MetTyrPheSerSer 254
Db 3124 GGGGTCTATTTTTGTGACACCTTTACTTGGCTTCCCTCGTTTGCTTATTTGGGCCAA 3183
Qy 255 ValValProValIleuGlnLysValAsnSerThrlYrThrlYsProValIleuAlaTProTyr 274
Db 3184 GGGACCCCGGTCAACGCTCCAGTGTGTAAGCCACCAAGCGCAGCGCCGACCAACCA 3243
Qy 275 SerProValHisProThrGlyThrSerGlnPro-----GlnArgProGluAspCysArg 292
Db 3244 ACAACCGCG---CCACCATCGCTGCGAGCCCTGTCTCCCTGCGCCAGAGCGGCGCTCG 3300
Qy 293 ProArg-----GlySerValLysGlyThrlYrThrlYsAspPheLeuGlnAspProLysIleu 310
Db 3301 CAGCGCGCGGGGGCGAGTGCACAGAGGGGCTGACTTGCCCTGAGTCCCAACTC 3360
Qy 311 CysTyrIleuLeuAspGlyIleLeuPheIleTyrGlyValIleIleThrlaLeuTyrIleu 330
Db 3361 TGCTACCTGCTGAGTAATCTCTTCACTATGATGATCTATCTCATCTGCTGCTTGTCCG 3420
Qy 331 ArgAlaIlePheSerArgSerAlaGlnThrAlaAlaAsnLeuGlnAspProAsnGlnLeu 350
Db 3421 AGAGTGAAGTTCAGCAGAGCGCAGAGCCCGCCGTACAGAGCGCAGAACCGAGTTC 3480
Qy 351 TyrAsnGlnLeuAsnLeuGlnArgArgGlnGluTyrTyrAspValIleuGlnLysIleArgAla 370
Db 3481 TATTAAGAGCTCAATCTTAGAGCAAGAGAGAGTACGATCTTTTGGACAAAGACGCTGGC 3540
Qy 371 ArgAspProGluMetGlyGlyLysGlnGlnArgArgAsnProGlnGlnValTyr 390
Db 3541 CGGAGCCTGAGATGGGGGAGAAAG---CCGAGAGAGAGAACCTTCAGAGAGGCTGTAC 3597
Qy 391 AsnAlaLeuGlnLysAspLysMetAlaGlnAlaTyrSerGlnIleGlyThrlYsGlyGln 410
Db 3598 AATGAACCTGAGAAAGATAGATGGCGAGGCTTACAGTGAAGATTGGATGAAAGCGAG 3657
Qy 411 ArgArgArgGlyLysGlyHisAspGlyLeuTyrGlnGlyLeuSerThrlaThrlYsAsp 430
Db 3658 CGCGGAGGGGCAAGGGGACCATGGCTTTACAGGGGTCTCAGTACAGCCACCAAGAGAC 3717
Qy 431 ThrTyrAspAlaLeuHlSwetGlnThrlLeuAlaProArg 443
Db 3718 ACCTACGACGCGCTTCACTAGAGGCGCTCCGCCCTCGC 3756

RESULT 4
US-10-006-771A-1
; Sequence 1, Application US/10006771A
; Publication No. US20020165360A1
; GENERAL INFORMATION:
; APPLICANT: Junghans, Richard P.
; TITLE OF INVENTION: Chimeric Effector Cell Receptors Against Carcinoembryonic Antigen
; FILE REFERENCE: 002
; CURRENT APPLICATION NUMBER: US/10/006,771A
; CURRENT FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: 60/250,090
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 7654
; TYPE: DNA

```



```

Db 1 CAGGTGAAGCTGCAGACGATCCGAGTTGAAGAACCGGGGTGAGACCTCAAGATC 60
Qy 21 SerCySLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValIysGlnAla 40
Db 61 AGCTGCAAGGCTTCGTGGTATACCTTCAACCGGTGGTATGAACCTGGGTGAAGCAGGCT 120
Qy 41 ProGlyGlnGlyLeuLysTyrMetGlyTyrPheLeuAsnThrSerThrGlyGlnSerThrPhe 60
Db 121 CCGGGCAAGGCTTTAAAGTGAAGTGGCTGGATTAAACCAAAACCTGGTGAAGCAACCTAT 180
Qy 61 AlaAspAspPheLysGlyArgPheAspPheSerLeuGlnThrSerAlaAsnThrAlaTyr 80
Db 181 GTTGAAGAGTTTAAAGGTGCTTTCCTCTCTTGGAGACCTCTGCCACCACTGCCAT 240
Qy 81 LeuGlnIleAsnAsnLeuLysSerGlnAspMetAlaThrTyrPheCysAlaArgTyrGln 100
Db 241 TTGCAGATCAACCACTCAAAATAGAGCACCGCTAAATATTTCTGTGCACGTTGGGAC 300
Qy 101 ValTyrHisGlyTyrVal-----ProTyrTyrPheGlyGlnGlyThrValThrVal 117
Db 301 TTCTAT---GATTACGTGAAGCTATGGAATTACTGGGGCCAAAGGACCAAGCTACCGTTC 357
Qy 118 Ser-----SerGlyGlyGlySerGly 125
Db 358 TCCAAGATCTCTGGTGGCGGTGGCTCGGGCGGTGGTGGTGGCGGCGCTCGGT 417
Qy 126 GlyGlyGlySerGlyGlyGlyGly-----SerAspIleGlnLeuThrGlnSerHisLysPhe 144
Db 418 GGTGGTGGTGGTGGCGGGCGGGCTCGAGCGATATGTGATGACCCAGTCTCAACGTTTC 477
Qy 145 LeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSerGlnAspValTyr 164
Db 478 ATGTCCACTTACAGTAGAGTATCGTACGCTCAGCTCAAAAGCAGTCAAGATGGGT 537
Qy 165 AsnAlaValAlaTyrTyrGlnGlnLysProGlyGlnSerProLysLeuLeuIleTyrSer 184
Db 538 ACGAATGTTGCTGGTATCAACAGAAACCGGGTCAATCCCGAAAGCACTGATTACTCG 597
Qy 185 AlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGlySerGlyProAsp 204
Db 598 GCATCTCAACCTTACAGTGGTGGTCCCGATCGCTTCAACCGGCAAGTGTCTGGGACCGAT 657
Qy 205 PheThrPheThrIleSerSerValGlnAlaGlnAspLeuAlaValTyrPheCysGlnGln 224
Db 658 TTCAACGCTCAACCAATCAGCAATGATCAGTCAAGACCTTGGGCGAGTATTTCTGTCATCAA 717
Qy 225 HisPheArgThrPro---PheThrPheGlySerGlyThrLysLeuGlnIleLys 241
Db 718 TATTACACTATCCGTTATTCACTGGGCTCGGGGACCAAGTTGGAATGAAAG 771

```

RESULT 6

```

US-10-150-762-38
; Sequence 38, Application US/10150762
; Publication No. US20030103948A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen C.
; APPLICANT: Graves, Scott S.
; APPLICANT: Schultz, Joanne E.
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James A.
; APPLICANT: Reno, John M.
; APPLICANT: Dearstyne, Erica A.
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; FILE REFERENCE: 690022.547C2
; CURRENT APPLICATION NUMBER: US/10/150.762
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 771
; TYPE: DNA

```

```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PRIA3 single chain antibody-genomic Streptavidin
; OTHER INFORMATION: fusion construct
US-10-150-762-38

```

Alignment Scores:

```

Pred. No.: 2.04e-94 Length: 771
Score: 994.00 Matches: 193
Percent Similarity: 82.95% Conservative: 21
Best Local Similarity: 74.81% Mismatches: 26
Query Match: 42.39% Indels: 18
DB: 15 Gaps: 5

```

US-09-596-774-7 (1-443) x US-10-150-762-38 (1-771)

```

Qy 1 GlnIleGlnLeuValGlnSerGlyProGlnLeuLysProGlyGlnThrValIle 20
Db 1 CAGGTGAAGCTGCAGACGATCCGAGTTGAAGAACCGGGGTGAGACCTCAAGATC 60
Qy 21 SerCySLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValIysGlnAla 40
Db 61 AGCTGCAAGGCTTCGTGGTATACCTTCAACCGGTGGTATGAACCTGGGTGAAGCAGGCT 120
Qy 41 ProGlyGlnGlyLeuLysTyrMetGlyTyrPheLeuAsnThrSerThrGlyGlnSerThrPhe 60
Db 121 CCGGGCAAGGCTTTAAAGTGAAGTGGCTGGATTAAACCAAAACCTGGTGAAGCAACCTAT 180
Qy 61 AlaAspAspPheLysGlyArgPheAspPheSerLeuGlnThrSerAlaAsnThrAlaTyr 80
Db 181 GTTGAAGAGTTTAAAGGTGCTTTCCTCTCTTGGAGACCTCTGCCACCACTGCCAT 240
Qy 81 LeuGlnIleAsnAsnLeuLysSerGlnAspMetAlaThrTyrPheCysAlaArgTyrGln 100
Db 241 TTGCAGATCAACCACTCAAAATAGAGCACCGCTAAATATTTCTGTGCACGTTGGGAC 300
Qy 101 ValTyrHisGlyTyrVal-----ProTyrTyrPheGlyGlnGlyThrValThrVal 117
Db 301 TTCTAT---GATTACGTGAAGCTATGGAATTACTGGGGCCAAAGGACCAAGCTACCGTTC 357
Qy 118 Ser-----SerGlyGlyGlySerGly 125
Db 358 TCCAAGATCTCTGGTGGCGGTGGCTCGGGCGGTGGTGGTGGCGGCGCTCGGT 417
Qy 126 GlyGlyGlySerGlyGlyGlyGly-----SerAspIleGlnLeuThrGlnSerHisLysPhe 144
Db 418 GGTGGTGGTGGTGGCGGGCGGGCTCGAGCGATATGTGATGACCCAGTCTCAACGTTTC 477
Qy 145 LeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSerGlnAspValTyr 164
Db 478 ATGTCCACTTACAGTAGAGTATCGTACGCTCAGCTCAAAAGCAGTCAAGATGGGT 537
Qy 165 AsnAlaValAlaTyrTyrGlnGlnLysProGlyGlnSerProLysLeuLeuIleTyrSer 184
Db 538 ACGAATGTTGCTGGTATCAACAGAAACCGGGTCAATCCCGAAAGCACTGATTACTCG 597
Qy 185 AlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGlySerGlyProAsp 204
Db 598 GCATCTCAACCTTACAGTGGTGGTCCCGATCGCTTCAACCGGCAAGTGTCTGGGACCGAT 657
Qy 205 PheThrPheThrIleSerSerValGlnAlaGlnAspLeuAlaValTyrPheCysGlnGln 224
Db 658 TTCAACGCTCAACCAATCAGCAATGATCAGTCAAGACCTTGGGCGAGTATTTCTGTCATCAA 717
Qy 225 HisPheArgThrPro---PheThrPheGlySerGlyThrLysLeuGlnIleLys 241
Db 718 TATTACACTATCCGTTATTCACTGGGCTCGGGGACCAAGTTGGAATGAAAG 771

```

RESULT 7

```

US-10-244-821-38
; Sequence 38, Application US/10244821
; Publication No. US2003014323A1
; GENERAL INFORMATION:

```



```

US-09-596-774-7 (1-443) x US-09-887-853-1 (1-909)
QY 1 GlnIleGlnLeuValGlnSerGlyProGluLeuLysPheGlyGluThrValLysIle 20
Db 9 GAGATCCAAATGGTTCAGTCTGACCTGAGCTGAAGAGAGCTCGAGAGACAGTCAAGATC 68
QY 21 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValLysGlnAla 40
Db 69 TCCTGCAAGGCTTCTGGGTATACCTTCAACAACTATGAGATGAATGAGTGGGTGAGAGAGCT 128
QY 41 ProGlyGlnGlyLeuLysTyrMetGlyTyrTrpIleAsnThrSerThrGlyLysSerThrPhe 60
Db 129 CAGGAAGAGGTTTAAAGTGATGGCTGATGAACACCAACACCTGAGAGAGCCAAACATAT 188
QY 61 AlaAspAspPheLysGlyTyrPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80
Db 189 GCTGAAGAGTTCAAGAGAGAGGTTTCCCTTCTTGGAACTCTGCGCAGACAGTCCATAT 248
QY 81 LeuGlnIleAsnAsnLeuLysSerGlyLysAspMetAlaThrTyrPheCysAlaArgTyrGlu 100
Db 249 TTGCAGATCAACAACCTCAAAAATAGAGACAGGCTACATATTTCTGTGAAAGCAATTT 308
QY 101 ValTyrHisGlyTyrValProTyrTyrGlyGlnGlyThrValThrValSerSerGly 120
Db 309 ATTACCTAAGCGGGGCTTCTTAACCTGAGGAGGAGCTGCTGCTGCTGCTGCA--- 365
QY 121 GlyGlyLysSerGlyGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThrGln 140
Db 366 TCGAGCTCCTCGGATCTTCACTAGCGGTTCCAGCTCGAGCATATGTCATACCCAG 425
QY 141 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSer 160
Db 426 TCTCTTAATTCATGTCACGCTGAGTGAGAGACAGGCTCAGCATCTCTCGAAGGCCAGT 485
QY 161 GlnAspValTyrAsnAlaValAlaTyrTyrGlnGlnLysProGlyLysSerProLysLeu 180
Db 486 CAGAGTGTGACTGCTGCTGAGCTGTATCAACAAACAGGAGCAATCTCTTAACCTA 545
QY 181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200
Db 546 CTGATTTACTGACATCCACCCGAGACCTGATCTGCTTCAACAGGAGAGTGA 605
QY 201 SerCysProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 220
Db 606 TCTGGGACAGATTATCTTCAACCATCAGCAGTGTGAGGCTGAAGACTGGCACTTCAT 665
QY 221 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyThrLysLeuGluIle 240
Db 666 TACTGTACGACAACTTATAGAGTCCCTGACGTTCCGAGGGGAGACCAAGCTGAGATA 725
QY 241 Lys 241
Db 726 AAA 728

```

RESULT 9

```

US-10-683-547-1
; Sequence 1, Application US/10683547
; Publication No. US20050058638A1
; GENERAL INFORMATION:
; APPLICANT: Huston, J.
; APPLICANT: Houston, L.L.
; APPLICANT: Ring, D.
; APPLICANT: Opermann, H.
; TITLE OF INVENTION: BIOSYNTHETIC BINDING PROTEINS FOR IMMUNO-TARGETING
; FILE REFERENCE: CIBT-P01-130
; CURRENT APPLICATION NUMBER: US/10/683,547
; PRIOR FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: US/09/558,741
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 07/831,967
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1

```

```

; SEQ ID NO 1
; LENGTH: 909
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: 741P8 sFv'
; NAME/KEY: CDS
; LOCATION: (3) -(752)
; OTHER INFORMATION:
US-10-683-547-1
Alignment Scores:
Pred. No.: 9,52e-92 Length: 909
Score: 969.50 Matches: 183
Percent Similarity: 85.06% Conservative: 22
Best Local Similarity: 75.93% Mismatches: 35
Query Match: 41.34% Indels: 1
DB: 19 Gaps: 1
US-09-596-774-7 (1-443) x US-10-683-547-1 (1-909)
QY 1 GlnIleGlnLeuValGlnSerGlyProGluLeuLysPheGlyGluThrValLysIle 20
Db 9 GAGATCCAAATGGTTCAGTCTGACCTGAGCTGAAGAGAGCTCGAGAGACAGTCAAGATC 68
QY 21 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValLysGlnAla 40
Db 69 TCCTGCAAGGCTTCTGGGTATACCTTCAACAACTATGAGATGAATGAGTGGGTGAGAGAGCT 128
QY 41 ProGlyGlnGlyLeuLysTyrMetGlyTyrTrpIleAsnThrSerThrGlyLysSerThrPhe 60
Db 129 CAGGAAGAGGTTTAAAGTGATGGCTGATGAACACCAACACCTGAGAGAGCCAAACATAT 188
QY 61 AlaAspAspPheLysGlyTyrPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80
Db 189 GCTGAAGAGTTCAAGAGAGAGGTTTCCCTTCTTGGAACTCTGCGCAGACAGTCCATAT 248
QY 81 LeuGlnIleAsnAsnLeuLysSerGlyLysAspMetAlaThrTyrPheCysAlaArgTyrGlu 100
Db 426 TCTCTTAATTCATGTCACGCTGAGTGAGAGACAGGCTCAGCATCTCTCGAAGGCCAGT 485
QY 101 ValTyrHisGlyTyrValProTyrTyrGlyGlnGlyThrValThrValSerSerGly 120
Db 309 ATTACCTAAGCGGGGCTTCTTAACCTGAGGAGGAGCTGCTGCTGCTGCA--- 365
QY 121 GlyGlyLysSerGlyGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThrGln 140
Db 366 TCGAGCTCCTCGGATCTTCACTAGCGGTTCCAGCTCGAGCATATGTCATACCCAG 425
QY 141 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSer 160
Db 426 TCTCTTAATTCATGTCACGCTGAGTGAGAGACAGGCTCAGCATCTCTCGAAGGCCAGT 485
QY 161 GlnAspValTyrAsnAlaValAlaTyrTyrGlnGlnLysProGlyLysSerProLysLeu 180
Db 486 CAGAGTGTGACTGCTGCTGAGCTGTATCAACAAACAGGAGCAATCTCTTAACCTA 545
QY 181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200
Db 546 CTGATTTACTGACATCCACCCGAGACCTGATCTGCTTCAACAGGAGAGTGA 605
QY 201 SerCysProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 220
Db 606 TCTGGGACAGATTATCTTCAACCATCAGCAGTGTGAGGCTGAAGACTGGCACTTCAT 665
QY 221 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyThrLysLeuGluIle 240
Db 666 TACTGTACGACAACTTATAGAGTCCCTGACGTTCCGAGGGGAGACCAAGCTGAGATA 725
QY 241 Lys 241
Db 726 AAA 728

```

RESULT 10

US-08-812-393A-1

Sequence 1, Application US/08812393A

Publication No. US2001007152A1

GENERAL INFORMATION:

APPLICANT: SHERMAN, Linda A.

APPLICANT: LUSTGARTEN, Joseph

TITLE OF INVENTION: RECOMBINANT CONSTRUCTS ENCODING

TITLE OF INVENTION: T CELL RECEPTORS SPECIFIC FOR HUMAN HLA-RESTRICTED TUMOR

NUMBER OF SEQUENCES: 64

CORRESPONDENCE ADDRESSES:

ADDRESSEE: MORRISON & FOERSTER

STREET: 2000 Pennsylvania Avenue, NW, suite 5500

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20006-1888

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/812.393A

FILING DATE: 05-MAR-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Murashige, Kate H

REGISTRATION NUMBER: 29,959

REFERENCE/DOCKET NUMBER: 31333-20001.00

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-887-1500

TELEFAX: 202-822-0168

TELEX:

INFORMATION FOR SEQ ID NO. 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1350 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

FEATURE:

NAME/KEY: Coding Sequence

LOCATION: 1...1332

OTHER INFORMATION:

US-08-812-393A-1

Alignment Scores:

Pred. No.: 2,28e-87

Score: 930.00

Percent Similarity: 59.33%

Best Local Similarity: 48.22%

Query Match: 39.66%

DB: 8

Gaps: 14

US-09-596-774-7 (1-443) x US-08-812-393A-1 (1-1350)

QY 3 GlnleuValGlnSerGlyProGlu-----LeuIysLysProGlyGluThrValLysIle 20
DB 70 CAGGAAGGCGAGCAGAGTCCCGCATCTTGATTCGCGAGGAGGAGACGACGAGCTC 129
QY 21 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValLysGlnAla 40
DB 130 CAGGTGAGCTTTTC-----ATCTTACAAACG---GTGCACTGCTTTTACCAACT 180
QY 41 ProGlyGlnGlyLeuLysTyrMetGlyTyrIleAsnThrSerThrGlyGluSerThrPhe 60
DB 181 CCGGAGGAGAGACGTCGTCAGCTGTGTGACATCTTCTCGGACAAAGCAGAGT----- 234

QY 61 AlaAspAspPheLysGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80
DB 235 -----GGGAGAGCTGACATCCACAGATCATTAAGAACGTCGAGCTCT 279
QY 81 LeuGlnIleAsnAsnLeuLysSerGluAspMetAlaThrTyrPheCysAlaArgTyrGlu 100
DB 280 TTGCACATTTCTCTCTCCCATCATCAGACTCAGACGACTTATCTCTGCTCAAAATCT 339
QY 101 ValTyrHisGlyTyrValProTyrTyrGlnGlyThrThrValThrValSerSerGly 120
DB 340 GGAGAAAGCAATGCAAGCTAACCTTACCTTGGGAAAGCACTAACTCTGTTAAATCACT 399
QY 121 GlyGlyLysSerGlyGlyGlyGlySerGlyGlyGlyLysSerAspIleGlnLeuThrGln 140
DB 400 GCGGAGAGGCTGTGCGGGGGTGGATCCGGGGGTGAGAGCTCAGAGGCTGACGCCACCA 459
QY 141 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSer 160
DB 460 AGCCCAAGAAACAAAGCTGCGAGTAACAGAGAAAGGTGACATTGAGCTGTATCAGACT 519
QY 161 GlnAspValTyrAsnAlaValAlaTyrTyrGlnGlnLysProGlyGlnSerProLysLeu 180
DB 520 AATAAC---CACAACAATGTAATCTGATGCGAGACAGCGGAGCATGGCTGAGCTG 576
QY 181 LeuIleTyrSer-----AlaSerSerArgTyrThrGlyValProSerArgPheThr 197
DB 577 ATCCATTTATCATATGCTGCGGACGACGACGACGACGACGACGACGACGACGACGAC 636
QY 198 GlySer---GlySerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAsp 216
DB 637 GCCTCCAGACCAACCAAGAGAACTCTCCCTCATCTGAGATTGAGTACCCCTCAG 696
QY 217 LeuAlaValTyrPheCys-----GlnGlnHisPheAsnThrProPheThrPheGly 233
DB 697 ACATCAGGTACTCTTGCCAGCGGTGACAGACGACCAACGAAAGATTAATTTTGCT 756
QY 234 SerGlyThrLysLeuGlnIleLysAlaLeuGlnIleSerAsnSerValMetTyrPheSer 253
DB 757 CATGGAACCAAGCTGTCTGTCTGACT-----AGTAATCTCATCATGATTAATTCAGC 807
QY 254 SerValValProValLeuGlnLysValAsnSerThrThrLysProValLeuArgThr 273
DB 808 CACTTCGTCGCGCTCTCTCCAGCGAGCCCAACGACG---CCAGCGCCCGGACCA 864
QY 274 ProSerProValHisProThrGlnGlyThrSerGlnProGlnArgProGluAspCysArgPro 293
DB 865 CCACACCGCGG---CCACCATCGCGTCGAGGCCCTG-----TCCCTGCGGCCCA 912
QY 294 ArgGlySerValLysGlyThrGlyLeuAspPheLeuGlnAspProLysLeuCysTyrLeu 313
DB 913 TCTAGTCT-----AGAGATCCCAACTGCTTACTCT 945
QY 314 LeuAspGlyIleLeuPheIleTyrGlyValIleIleThrAlaLeuTyrLeuArgAlaLys 333
DB 946 CTGATGGAATCTCTCATCTATGATGATCTCATCTGCTGCTGCTGCTGCTGCTGCTGCT 1005
QY 334 PheSerArgSerAlaGluThrAlaAlaAsnLeuGlnAspProAsnGlnIleTyrAsnGlu 353
DB 1006 TTCAGCAGAGCGGACGCGCCCGCTTACAGAGGCGGACGAAACGAGCTTAAACAGAG 1065
QY 354 LeuAsnLeuGlyArgGluGluTyrAspValLeuGlnLysLysAlaArgAspPro 373
DB 1066 CTCAATCTTAGACCAAGAGAGAGTACGATTTTGGACAAAGACGTCGCGGAGACCT 1125
QY 374 GluMetGlyGlyLysGlnGlnArgArgAsnProGlnGlnGlyValTyrAsnAlaLeu 393
DB 1126 GAGATGGGGGAAAG---CCGAGAAAGAAAGACCTCAGAGAGGCTGTAACAAAGAACTG 1182
QY 394 GlnLysAspLysMetAlaGluAlaTyrSerGlnIleLysThrLysGlyGluArgArgArg 413
DB 1183 CAGAAATATAGATGGGAGGCGCTTACAGTGAATTTGGATGAAGGCGGAGCGGAGG 1242
QY 414 GlyLysGlyHisAspGlyLeuTyrGlnGlyLeuSerThrAlaThrLysAspThrTyrAsp 433


```

Db      1243  GGCAAGGGGACGATGGCTTTACCAAGGCTCTCAATGACGACCAAGAGACCTTACGAC 1302
Qy      434  AAlaLeuHleuMeGlnThrLeuAlaProArg 443
Db      1303  GCCCTTCAATGACGAGGCCCTGCCCCCTCGC 1332

RESULT 11
US-09-774-681-1
; Sequence 1, Application US/09774681
; Publication No. US20030208780A1
; GENERAL INFORMATION:
; APPLICANT: Sunol Molecular Corporation
; APPLICANT: Sherman, Linda
; APPLICANT: Lustgarten, Joseph
; TITLE OF INVENTION: RECOMBINANT CONSTRUCTS ENCODING T CELL
; FILE REFERENCE: 31333-20001.01
; CURRENT APPLICATION NUMBER: US/09/774,681
; PRIOR FILING DATE: 2001-02-01
; PRIOR APPLICATION NUMBER: US 08/812,393
; PRIOR FILING DATE: 1997-03-05
; PRIOR APPLICATION NUMBER: US 60/012,845
; NUMBER OF SEQ ID NOS: 65
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1350
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Polynucleotide derivative of effective T cell
; NAME/KEY: CDS
; LOCATION: (1)...(1350)
US-09-774-681-1

Alignment Scores:
Pred. No.:      3.7e-87      Length:      1350
Score:          928.00      Matches:      217
Percent Similarity: 59.33%      Conservative: 50
Best Local Similarity: 48.22%      Mismatches: 145
Query Match:      39.57%      Indels:      38
DB:               10          Gaps:      14

US-09-596-774-7 (1-443) x US-09-774-681-1 (1-1350)
Qy      3  GlnLeuValGlnSerGlyProGlu-----LeuLysLysProGlyGlnThrValLysIle 20
Db      70  CAGCAAGTGCAGCAGAGTCCCGCATCTCTGCTTCGACGAGGAGGAGAAACGACAGCTC 129
Qy      21  SerCysLysAlaSerGlyTyrProPheThrAspTyrGlyMetAsnThrPvalLysGlnAla 40
Db      130  CAGGTGAGCTTTCC-----ATCTTTACAAACAG--GTGCGGTGTTTACCAACGT 180
Qy      41  ProGlyGlnGlyLeuLysTyrMetGlyTyrPheAsnThrSerThrGlyLysSerThrPhe 60
Db      181  CCTGGGGAGAGACTGTCAGCGCTTGTGACAAATCTCTGGGACAAAGCAGAGT----- 234
Qy      61  AlaAspAspPheLysGlyTyrGpPheAspPheSerLeuGlnThrSerAlaAsnThrAlaTyr 80
Db      235  -----GGGAGACTGACATCCACCAACAGTCAATTAAGAACCTCCGAGCTT 279
Qy      81  LeuGlnIleAsnLeuLysSerGlyAspMetAlaThrTyrPheCysAlaArgTyrPglu 100
Db      280  TTGACACATTTCTCTCTCCGATCAGACTCCAGGCACTTAATCTCTGTGCTCAAAATCT 339
Qy      101  ValTyrHisGlyTyrValProTyrTyrPglGlnGlyThrThrValThrValSerSerGly 120
Db      340  GGAAGAACATGCAAGAGTAACTTACCTGGGAAAGGCACTAACTCTCTGTAAATACAGT 399
Qy      121  GlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlyGlyGlyGly 140

```

```

Db      400  GGCGAGGCTGTGGCGGGGAGATCCGAGGAGTGAAGCTCAGAGCTCGATACCCAA 459
Qy      141  SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSer 160
Db      460  AGCCCAAGAAACAAAGTGGCAGTAAACAGAGAAAGTGCATTAAGCTCTAATACACT 519
Qy      161  GlnAspValTyrAsnAlaValAlaTyrPglGlnGlnLysProGlyGlnSerProLysLeu 180
Db      520  AATAAAC---CACCAACAAATGACTGTATCTGGCAGACAGGAGGAGTGGAGGCTG 576
Qy      181  LeuIleTyrSer-----AlaSerSerArgTyrThrGlyValProSerArgPheThr 197
Db      577  ATCATTAATTCATATGCTGTGGAGAGACTGAGAAAGAAATATTCCTGATGATACAG 636
Qy      198  GlySer---GlySerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAsp 216
Db      637  GCCCTCAGACCAAGCCAAAGAAACTCTCCCTCATTTGTGAGTGGTATCCCTCTCAG 696
Qy      217  LeuAlaValTyrPheCys-----GlnGlnHisPheArgThrProPheThrPheGly 233
Db      697  ACATCACTGTACTTCTGTGCGCAGCGGTGAGACAGGACCAACAAAGATTATTTTCGT 756
Qy      234  SerGlyThrLysLeuGlnIleLysAlaLeuGlnIleSerAsnSerValMetTyrPheSer 253
Db      757  CATGGAACCAAGCTGTCTGTCTGACT-----AGTAATCTCATCATGATCTTCTCAG 807
Qy      254  SerValValProValLeuGlnLysValAsnSerThrThrLysProValLeuArgThr 273
Db      808  CACTTCGTGGCGGTCTCTCTGCGAGGAAAGCCACACGACG---CCAGGCGCGCAGCA 864
Qy      274  ProSerProValHisProThrGlyThrSerGlnProGlnArgProGlnAspCysArgPro 293
Db      865  CCAACACCGCG---CCACCAACGCGCTCGACACCCCTG-----TCCCTGCGCCCA 912
Qy      294  ArgGlySerValLysGlyThrGlyLeuAspPheLeuGlnAspProLysLeuGlyTyrLeu 313
Db      913  TCTAGTTCT-----AGAGATCCCAACTCTGTCTACCTG 945
Qy      314  LeuAspGlyLysLeuPheIleTyrGlyValIleIleThrAlaLeuTyrLeuArgAlaLys 333
Db      946  CTGATGAGATCTCTTCAATCATATGATGTCATTCATTCAGCTCTGTCGAGAGTGAAG 1005
Qy      334  PheSerArgSerAlaGlnThrAlaAlaAsnLeuGlnAspProAsnGlnLeuTyrAsnGlu 353
Db      1006  TTCAGACAGAGCGCCACACCCCGCGTACACAGGCGCCAGAACCTCTATTAACGAG 1065
Qy      354  LeuAsnLeuGlyArgArgGlnGlnTyrAspValLeuGlnLysArgAlaArgAspPro 373
Db      1066  CTCATATCTAGAGCAAGAGAGAGTACGATGTTTGGACAAAGAGCTGGCGGAGCCT 1125
Qy      374  GluMetGlyGlyLysGlnGlnArgArgAsnProGlnGlnGlyValTyrAsnAlaLeu 393
Db      1126  GAGATGGGGGAGAAAG---CCGAGAGAGAAAGAACTCCAGAGAGGCTGTACATGAATG 1182
Qy      394  GlnLysAspLysMetAlaGlnAlaTyrSerGlnIleGlyThrLysGlyValLysArgArg 413
Db      1183  CAGAAAGATTAAGTGGCGGAGGCTTACATGATGATGGATGAAGAGCAGCGCGGAGG 1242
Qy      414  GlyLysGlyHisAspGlyLeuTyrGlnGlyLysSerThrAlaThrLysAspThrTyrAsp 433
Db      1243  GGCAAGGGGACAGATGGCTTTACAGAGGATCTCAGTACAGCAGCAGCAAGGACCTACGAC 1302
Qy      434  AAlaLeuHleuMeGlnThrLeuAlaProArg 443
Db      1303  GCCCTTCAATGACGAGGCCCTGCCCCCTCGC 1332

RESULT 12
US-10-239-656-78
; Sequence 78, Application US/10239656
; Publication No. US20040038339A1
; GENERAL INFORMATION:
; APPLICANT: KUFER, PETER
; APPLICANT: RIETHMULLER, GERT

```


APPLICANT: LUTTERBUSE, RALF
 APPLICANT: BORSCHERT, KATRIN
 APPLICANT: KISCHEL, ROMAN
 APPLICANT: MAYER, MONIKA
 APPLICANT: HOFMEISTER, ROBERT
 TITLE OF INVENTION: MULTIFUNCTIONAL POLYPEPTIDES COMPRISING A BINDING SITE
 FILE REFERENCE: TO AN EPITOP OF THE NKGD RECEPTOR COMPLEX
 CURRENT FILING DATE: 2003-03-06
 PRIOR APPLICATION NUMBER: US/10/239,656
 PRIOR FILING DATE: 2001-03-26
 PRIOR APPLICATION NUMBER: EP 00106467.4
 PRIOR FILING DATE: 2000-03-24
 NUMBER OF SEQ ID NOS: 92
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO: 78
 LENGTH: 1515
 TYPE: DNA
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Description of Artificial Sequence: Synthetic 3810xps-
 US-10-239-656-78

Alignment Scores:

Pred. No.: 5,476-86 Length: 1515
 Score: 917.50 Matches: 185
 Percent Similarity: 57.38% Conservative: 25
 Best Local Similarity: 50.55% Mismatches: 31
 Query Match: 39.13% Indels: 125
 Gaps: 4

US-09-596-774-7 (1-443) x US-10-239-656-78 (1-1515)

QY 1 GlnIleGlnLeuValGlnSerGlyProGluLeuLysLysProGluGluThrValLysIle 20
 DB 382 CAGGTGCAACTGCAGCAGCTGAGGCGCTGAGTGAAGAAAGCCGAGAGAGAGAGTCAAGATC 441
 QY 21 SerCysLysAlaSerGlyTyrProPheThrIleValMetLeuTyrValLysGlnAla 40
 DB 442 TCCTGCAAGGCTTCGGGTATACCTTCACAACTGTGAATGAACCTGGGTGAAGCAGGCT 501
 QY 41 ProGluGlnGlyLeuValTyrMetGlyTyrIleAsnThrSerThrGlyGluSerThrPhe 60
 DB 502 CCAGGAAGGGGTTCAAGTGGAGTGGCTGATTAACCTTACACTGAGAGCCAAACATATC 561
 QY 61 AlaAspAspPheLysGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80
 DB 562 GGTATGACTTCAGAGGAGCGGTTGCCCTTCTTGAAACCTGCGCAGCAGCTGCCCTAT 621
 QY 81 LeuGlnIleAsnAsnLeuLysSerGluAspMetAlaThrTyrPheCysAlaArg----- 98
 DB 622 TTGCAGATCAACAACCTCAAAAATGAGACACGCTACATATTTCTGTCAAGATTCAAC 681
 QY 98 ----- 98
 DB 682 TCCCTGACTACTGAGGCGCAAGGACACGCTCACCGTCTCTCCGAGGTGGTGGATCC 741
 QY 98 ----- 98
 DB 742 GAGGTGCAAGCTCTGAGTCTGAGAGTGGCTGTGTCAGCCTCGAGAGATCCTTGAATCTC 801
 QY 98 ----- 98
 DB 802 TCCTGTGAGCCTCAGATTGATTTAGTAGATAGTAGATGATGGGTCCGAGAGCT 861
 QY 99 ----- 99
 DB 862 CCAGGAAAGGGCTTGAATGATTGGAGAAATTAATCCAGATAGCAGTAACTAAT 921
 QY 99 ----- 99

DB 922 AGCCATCTTAAGATAGATTCATCTCCAGAGCAACGCCAAATAACGCTGAC 981
 QY 100 -----GluValTyrHis----- 103
 DB 982 CTGCAATAGCAAAAGTAGAGCTGAGACACAGCCCTTTATTACTGTGCAAGATTGGG 1041
 QY 104 -----GlyTyrValProTyrTyrPdlGlnGlyThrThrValThrValSerSerGly 121
 DB 1042 CAATGGGGTACTTGTGACTACTGGGGCCAGAGGACACGCTGCTCTCCAGTGGT 1101
 QY 122 GlnGlySerGlyGlyGlyGlySerGlyGlyGlyGlySerAspIleGlnLeuThrGlnSer 141
 DB 1102 GGTGGTCTGGCGCGCGGCTCCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1161
 QY 142 HisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysValAlaSerGln 161
 DB 1162 CCATCTCTCTGACTGTGACGACAGAGAGGGTCACTATGAGCTGCAAGTCCAGTACG 1221
 QY 162 AspValTyrAsnAla-----ValAlaTyrGlnGlnLysProGly 175
 DB 1222 AGTCTGTTAAACAGTGAATCAAAAGAACTACTGACTGTGACACAGAGAAACAGAG 1281
 QY 176 GlnSerProLysLeuLeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArg 195
 DB 1282 CAGCCTCTTAAACTGTGTATCTACTGGGCAATCCACTAGGAATGTGGGGTCTCATGTCG 1341
 QY 196 PheThrGlySerGlySerGlyProAspPheThrPheThrIleSerSerValGlnAlaGlu 215
 DB 1342 TTCACAGGCACTGTGAGTGAACAGATTCACTCTCACATCAGACAGTGTGAGGCTGAA 1401
 QY 216 AspleuAlaValTyrPheCysGlnGlnHisPheArgThrProPheThrPheGlySerGly 235
 DB 1402 GACCTGCGAGTTATTACTGTGACAGATGATATATTATCTCTCACGCTTGGGTCTGG 1461
 QY 236 ThrLysLeuGlnIleLys 241
 DB 1462 ACCAAGCTTGAGATCAAA 1479

RESULT 13

US-09-766-543-9
 ; Sequence 9, Application US/09766543
 ; Patent No. US20020041865A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Austin, Richard
 ; APPLICANT: Kwok, Cheuk S.
 ; TITLE OF INVENTION: METHODS FOR TREATING TUMORS
 ; FILE REFERENCE: PP01679.002
 ; CURRENT APPLICATION NUMBER: US/09/766,543
 ; CURRENT FILING DATE: 2000-01-19
 ; PRIOR APPLICATION NUMBER: 60/117,258
 ; PRIOR FILING DATE: 2000-01-20
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 9
 ; LENGTH: 816
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence: 520C9
 ; OTHER INFORMATION: humanized single-chain antibody used in the
 ; OTHER INFORMATION: IL-2-antibody fusions
 ; NAME/KEY: CDS
 ; LOCATION: (7) ..(807)
 US-09-766-543-9

Alignment Scores:

Pred. No.: 7,036-86 Length: 816
 Score: 913.00 Matches: 167
 Percent Similarity: 83.61% Conservative: 37
 Best Local Similarity: 68.44% Mismatches: 36
 Query Match: 38.93% Indels: 4
 Gaps: 2

US-09-596-774-7 (1-443) x US-09-766-543-9 (1-816)

```

QY 1 GlnllleGlnleuValGlnSerGlyProGluLeuLysProGlyGlnThrValLysIle 20
Db 73 GAGATGCAACCTGGTGGAGCTGGGCTGAGGGAAGCCCTGGGCTCAGTGAAGCTC 132
QY 21 SerCyLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTrpValLysGln 40
Db 133 TCCTGCAAGGCTTCGTGTTACCTTTACCACTTGAATGAACTGGGTCGAGACAGC 192
QY 41 ProGlyGlnGlyLeuLysTrpMetGlyTyrIleAsnThrSerThrGlyLysSerThrPhe 60
Db 193 CCTGAGCAAGGCTTGAATGATGGATGATGATGAACACTACACTGACAGTCAACATAT 252
QY 61 AlaAspAspPheLysGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80
Db 253 GCTGATGACTTCAGAGAAAGAGTCCATGACACAGACATCCACAGACAGCCTTAC 312
QY 81 LeuGlnIleAsnAsnLeuLysSerGlyLysAspMetAlaThrTyrPheCysAlaArgTrpGlu 100
Db 313 ATGACACCTGAGAGAGCTGATGATCTGACAGACAGGCTGTATTTACTGTGCGAGA----- 366
QY 101 ValTyrHisGlyTyrValProTyrTrpGlyGlnGlyThrValThrValSerSerGly 120
Db 367 ---CGATTGGGATTGGCT---TACTGGGGCCAGGGAACCTGTGTCACCGTCTCCTCAGGT 420
QY 121 GlyGlyGlySerGlyGlyGlySerGlyGlyGlySerGlyGlySerAspIleGlnLeuThrGln 140
Db 421 GGGGTGGCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 480
QY 141 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSer 160
Db 481 TCTCATCTCCTCCCTGCTGCTGATCTGTAGGAGACAGAGTCACTGCGGGGCAAGT 540
QY 161 GlnAspValTyrAsnAlaValAlaTrpTyrGlnGlnLysProGlyGlnSerProLysLeu 180
Db 541 CAGACACATGGTAATAGCTTAACCTGTATCAGACAGAAACAGGAGAAACCCCTAAGCTC 600
QY 181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200
Db 601 CTGATCTACAGCCATCCAGCTTGAATTCGGGGTCCCATCAGTTCAGTGAAGTGA 660
QY 201 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 220
Db 661 TCTGGAGCAAGATTTTACTTTCACATCAGCAGCTGCGAGCTGGAATTTGCAACATAT 720
QY 221 PheCysGlnGlnHisPheArgTyrProPheThrPheGlySerGlyThrLysLeuGluIle 240
Db 721 TACGTCTACAAATATGCTATTTTTCGTACACGTTCCGCCAAGGAGACAGCTGAGATT 780
QY 241 LysAlaLeuGlu 244
Db 781 AAGGATCCGAA 792

```

RESULT 14
US-10-127-890-90

Sequence 90, Application US/10127890
Publication No. US20030166196A1

GENERAL INFORMATION:
APPLICANT: Better, Marc D.
Carroll, Stephen F.

Studilka, Gary M.
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating

Proteins
NUMBER OF SEQUENCES: 173

CORRESPONDENCE ADDRESS:
ADDRESSEE: McAndrews, Held & Malloy, Ltd.
STREET: 500 West Madison Street, 34th floor

CITY: Chicago
STATE: Illinois
COUNTRY: USA

ZIP: 60661

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/127,890

FILING DATE: 23-Apr-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/646,360

FILING DATE: 13-MAY-1996

APPLICATION NUMBER: PCT/US94/05348

FILING DATE: 12-MAY-1994

APPLICATION NUMBER: US 08/064,691

FILING DATE: 12-MAY-1993

APPLICATION NUMBER: US 07/986,430

FILING DATE: 09-DEC-1992

APPLICATION NUMBER: US 07/901,707

FILING DATE: 19-JUN-1992

APPLICATION NUMBER: US 07/787,567

FILING DATE: 04-NOV-1991

ATTORNEY/AGENT INFORMATION:

NAME: McNicholas, Janet M.

REGISTRATION NUMBER: 32,918

REFERENCE/DOCKET NUMBER: 200-70.P4

TELEPHONE: 312/707-8889

TELEFAX: 312/707-9155

TELEX: 650 388-1248

INFORMATION FOR SEQ ID NO: 90:

SEQUENCE CHARACTERISTICS:

LENGTH: 723 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA

SEQUENCE DESCRIPTION: SEQ ID NO: 90:

US-10-127-890-90

US-09-596-774-7 (1-443) x US-10-127-890-90 (1-723)

```

QY 1 GlnllleGlnleuValGlnSerGlyProGluLeuLysProGlyGlnThrValLysIle 20
Db 1 GAGATCCAGTGGTGGAGTCTGGAGAGGCTGTGGTGAACCTGGAGGCTCGTCAAAATC 60
QY 21 SerCyLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTrpValLysGln 40
Db 61 TCCTGCGAGCTTCGTGATATACCTTCAAACTATGATGATGATGATGATGATGATGATGAT 120
QY 41 ProGlyGlnGlyLeuLysTrpMetGlyTyrIleAsnThrSerThrGlyLysSerThrPhe 60
Db 121 CCAGGAAAGGTTTGAAGTGAATGGCTGATGAACCAACCACTGAGAGCCAAATAT 180
QY 61 AlaAspAspPheLysGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80
Db 181 GCTGATCTTTCAGAGGAGCGTTTACCTTCTTTGGACGATTCCTTAAGAACCTGCTAT 240
QY 81 LeuGlnIleAsnAsnLeuLysSerGlyLysAspMetAlaThrTyrPheCysAlaArgTrpGlu 100
Db 241 TTAAGATTAAGAGCTGAGACCGACAGACAGGCTGTATTTCTTCAAGA---CGG 297
QY 101 ValTyrHisGlyTyrValProTyrTrpGlyGlnGlyThrValThrValSerSerGly 120
Db 298 GGTACAGACTGACTGATGATGTCTGGGGCCAAAGGAGCAACGCTCAGCTCCTCAGGT 357

```

QY 121 GYGLYGLYSERGLYGLYGLYSERGLYGLYGLYSERAPPILEGLEUTHRGIN 140
 DB 358 GGGCGATGATTTGGTGGAGTGGTCCGAGAGTGGAGATCTGACATCAAGTGAATCAG 417
 QY 141 SerHisysPheLeuSerThrSerValGlyAspArgValSerIleThrCysAlaSer 160
 DB 418 TCTCCATCTTCCCTGCTGTCATCTGTAGAGACAGAGTCACTATCATTGCGGGCGAGT 477
 QY 161 GlnAspValTyrAsnAlaValAlaIleTyrGlnGlnLysProGlyGlnSerProLysLeu 180
 DB 478 CAGACATTAATTAATTAAGTGTTCAGAGAAACAGGAAAGCTCTTAAGGCC 537
 QY 181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200
 DB 538 CTGATCTATCTGCAACAGATGGATCTGGGCTCCATCAAGCTTCACTGAGTGGAG 597
 QY 201 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 220
 DB 598 TCTGGACAGATTAATTAATCTCAACATCAGAGCTCGCAATTAAGATTTTGAATTAT 657
 QY 221 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyThrLysLeuGluIle 240
 DB 658 TATGTCAACAGTATGATGAGTCTCCGTGAGCGTTCGGTGGAGCACCAAGCTTGAGATG 717
 QY 241 Lys 241
 DB 718 AAA 720
 RESULT 15
 US-10-717-243-90
 ; Sequence 90. Application US/10717243
 ; Publication NO. US2005054835A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Better, Marc D.
 ; Caroli, Stephen F.
 ; Studinka, Gary M.
 ; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
 ; NUMBER OF SEQUENCES: 169
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: McAndrews, Held & Malloy, Ltd.
 ; STREET: 500 West Madison Street, 34th floor
 ; CITY: Chicago
 ; STATE: Illinois
 ; COUNTRY: USA
 ; ZIP: 60661
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/10/717,243
 ; FILING DATE: 18-Nov-2003
 ; CLASSIFICATION: 530
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/839,765
 ; FILING DATE: 15-APR-1997
 ; APPLICATION NUMBER: US 08/425,336
 ; FILING DATE: 18-APR-1995
 ; APPLICATION NUMBER: US 08/064,691
 ; FILING DATE: 12-MAY-1993
 ; APPLICATION NUMBER: US 07/988,430
 ; FILING DATE: 09-DEC-1992
 ; APPLICATION NUMBER: US 07/901,707
 ; FILING DATE: 19-JUN-1992
 ; APPLICATION NUMBER: US 07/787,567
 ; FILING DATE: 04-NOV-1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: McNicholas, Janet M.
 ; REGISTRATION NUMBER: 32,918
 ; REFERENCE/DOCKET NUMBER: 11022US09/200-70.P3.C3

; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 312/707-8889
 ; TELEFAX: 312/707-9155
 ; TELEX: 650 388-1248
 ; INFORMATION FOR SEQ ID NO: 90:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 723 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 90:
 ; US-10-717-243-90
 Alignment Scores:
 Pred. No.: 5,93e-85 Length: 723
 Score: 903.50 Matches: 167
 Percent Similarity: 82.57% Conservative: 32
 Best Local Similarity: 69.29% Mismatches: 41
 Query Match: 38.53% Indels: 1
 DB: 19 Gaps: 1
 US-09-596-774-7 (1-443) x US-10-717-243-90 (1-723)
 QY 1 GlnIleGlnLeuValGlnSerGlyProGluLeuLysPheProGlyGlnThrValLysIle 20
 DB 1 GAGATCCAGTTGGTGGAGTCTGGAGAGAGCGCTGGTGAAGCTTGAGAGGCTCGTCAATC 60
 QY 21 SerCysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValLysGlnAla 40
 DB 61 TCTGCCAGCTCTGGGTATATACCTTCAAACTATGAAATGAAGTGGTGGCCAGCT 120
 QY 41 ProGlyGlnGlyLeuLysTyrMetGlyTyrTrpIleAsnThrSerThrGlyLysSerThrPhe 60
 DB 121 CCAGAAAGGTTAGAGTGGATGGCTGATTAACCCACACTGGAGAGCCAACTAT 180
 QY 61 AlaAspAspPheLysGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80
 DB 181 GCTATCTCTTCAAGGAGACGTTTACCTTCTCTTGGACATTTAGAACACAGTCCAT 240
 QY 81 LeuGlnIleAsnLeuLysSerGlyAspMetAlaThrTyrPheCysAlaArgTyrGlu 100
 DB 241 TTACAGATCAACAGCCACAGCCGAGAGACACGCTGTGATTTCTGTACAAAGA---CGG 297
 QY 101 ValTyrHisGlyTyrValProTyrTyrGlnGlnLysThrValThrValSerSerGly 120
 DB 298 GGTACACACTGTACTTCGATGCTGTGGGCCCAAGGACCAAGTCAACCGTCTCTCAGGT 357
 QY 121 GYGLYGLYSERGLYGLYGLYSERGLYGLYGLYSERAPPILEGLEUTHRGIN 140
 DB 358 GGGCGATGATTTGGTGGAGTGGTCCGAGAGTGGAGATCTGACATCAAGTGAATCAG 417
 QY 141 SerHisysPheLeuSerThrSerValGlyAspArgValSerIleThrCysAlaSer 160
 DB 418 TCTCCATCTTCCCTGCTGTCATCTGTAGAGACAGAGTCACTATCATTGCGGGCGAGT 477
 QY 161 GlnAspValTyrAsnAlaValAlaIleTyrGlnGlnLysProGlyGlnSerProLysLeu 180
 DB 478 CAGACATTAATTAATTAAGTGTTCAGAGAAACAGGAAAGCTCTTAAGGCC 537
 QY 181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200
 DB 538 CTGATCTATCTGCAACAGATGGATCTGGGCTCCATCAAGCTTCACTGAGTGGAG 597
 QY 201 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 220
 DB 598 TCTGGACAGATTAATTAATCTCAACATCAGAGCTCGCAATTAAGATTTTGAATTAT 657
 QY 221 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyThrLysLeuGluIle 240
 DB 658 TATGTCAACAGTATGATGAGTCTCCGTGAGCGTTCGGTGGAGCACCAAGCTTGAGATG 717
 QY 241 Lys 241

Thu Apr 21 06:48:36 2005

us-09-596-774-7.apr20.rnpb

Page 14

Db : 718 AAA 720

Search completed: April 20, 2005, 17:07:55
Job time : 781.098 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2005, 06:53:52 ; Search time 17.1515 Seconds
(without alignments)
2485.142 Million cell updates/sec

Title: US-09-596-774-7
Perfect score: 2345
Sequence: 1 QIOLVQSGPELKKRGETVKI.....LSTATKDYDALHMOTLAPR 443

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR 79: *
1: p1r1: *
2: p1r2: *
3: p1r3: *
4: p1r4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	713	30.4	164	2	A40104 T-cell receptor CD
2	694	29.6	268	2	A56446 Ig heavy chain V r
3	665.5	28.4	249	2	S41374 single chain Fv an
4	611.5	26.1	233	2	JC5322 p53 specific singl
5	605	25.8	177	2	S54817 T-cell receptor io
6	605	25.8	206	2	A35900 T-cell receptor CD
7	604.5	25.8	163	2	A31768 T-cell receptor ze
8	589.5	25.1	188	2	A45089 T-cell receptor CD
9	539	23.0	166	2	JC4664 T-cell receptor ze
10	537.5	22.9	166	2	I46424 T-cell surface gly
11	537.5	22.9	120	2	B42848 L6 mab heavy chain
12	535.5	22.8	146	4	S32190 Ig heavy chain pre
13	525	22.4	136	2	S35759 BHD9D10 protein-
14	522.5	22.3	118	2	S37204 Ig heavy chain V r
15	519	22.1	119	2	A53285 Ig heavy chain V a
16	516	22.0	119	2	H45722 anti-glycoprotein
17	510	21.7	117	2	S32190 Ig heavy chain V r
18	499.5	21.3	115	2	S19665 Ig heavy chain V r
19	495	21.1	107	2	S32192 Ig kappa chain V r
20	490	20.9	107	2	S32191 Ig kappa chain V r
21	490	20.9	113	2	B36259 Ig heavy chain V r
22	488.5	20.8	118	2	A32530 Ig heavy chain V r
23	488.5	20.8	139	2	PH1225 Ig heavy chain pre
24	487	20.8	124	2	PH1404 Ig heavy chain V r
25	486.5	20.7	115	2	S19668 Ig heavy chain V r
26	486	20.7	117	2	S32187 Ig heavy chain V r
27	486	20.7	118	2	S19667 Ig heavy chain V r
28	486	20.7	149	1	KVMS11 Ig kappa chain pre
29	484.5	20.7	109	2	S26325 Ig heavy chain V r

30	484.5	20.7	120	2	S19663 Ig heavy chain V r
31	484	20.6	119	2	B32530 Ig heavy chain V r
32	480	20.5	119	2	P00265 Ig kappa chain V r
33	477.5	20.4	114	2	D32967 Ig heavy chain V r
34	472.5	20.1	114	2	C32967 Ig heavy chain V r
35	469	20.0	152	2	S30751 Ig kappa chain pre
36	460	19.6	114	2	PLO256 Ig heavy chain V r
37	453.5	19.3	108	2	PLO083 Ig kappa chain V r
38	452	19.3	117	2	S42466 Ig kappa chain V r
39	451	19.2	131	2	S26792 Ig heavy chain V r
40	450	19.2	105	2	S24765 Ig heavy chain V r
41	448.5	19.1	127	2	S04577 Ig kappa chain pre
42	446	19.0	131	2	PLO207 Ig heavy chain V r
43	443.5	18.9	101	2	D24672 Ig heavy chain V r
44	442	18.8	107	2	S09667 Ig kappa chain V-J
45	442	18.8	134	2	S21916 Ig heavy chain V r

ALIGNMENTS

RESULT 1
A40104
T-cell receptor CD3 zeta chain - mouse
C:Species: Mus musculus (house mouse)
C:Date: 20-Mar-1992 #sequence_revision 20-Mar-1992 #text_change 09-Jul-2004
C:Accession: A40104; I55293
R:Weisman, A.M.; Baniyash, M.; Hou, D.; Samelson, L.E.; Burgess, W.H.; Klausner, R.D.
A:Title: Molecular cloning of the zeta chain of the T cell antigen receptor.
A:Reference number: A40104; MUID:88145643; PMID:3278377
A:Accession: A40104
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-164 <WEI>
A:Cross-references: UNIPROT:P24161; GB:M19729; NID:G201131; PID:AAA40171.1; PID:G201132
R:Baniyash, M.; Heu, V.W.; Seldin, M.F.; Klausner, R.D.
J Biol Chem. 264, 13252-13257, 1989
A:Title: The isolation and characterization of the murine T cell antigen receptor zeta cl
A:Reference number: I55293; MUID:88327299; PMID:2787796
A:Accession: I55293
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-164 <RES>
A:Cross-references: GB:J04967; NID:G556326; PID:AAA50301.1; PID:G556327
A:Gene: Tcrz
C:Keywords: phosphoprotein; T-cell receptor; transmembrane protein
Query Match 30.4%; Score 713; DB 2; Length 164;
Best local similarity 99.3%; Pred. No. 5.4e-43; Indels 0; Gaps 0;
Matches 138; Conservative 0; Mismatches 1;
OY 305 LEDPKCYLIDGLIFIVGVIITLALYLRKFSRSATNANLQDPNOLYNELNGRREYDV 364
DB 26 LDDPKCYLIDGLIFIVGVIITLALYLRKFSRSATNANLQDPNOLYNELNGRREYDV 85
OY 365 LEKRRADPEWKGKQQRNRRPQEGVYNALQDKMAEYSEIGTGERRRGKHGGLYQGL 424
DB 86 LEKRRADPEWKGKQQRNRRPQEGVYNALQDKMAEYSEIGTGERRRGKHGGLYQGL 145
OY 425 STATKDYDALHMOTLAPR 443
DB 146 STATKDYDALHMOTLAPR 164
RESULT 2
A56446
Ig heavy chain V region (3H-3H scFv) - mouse (strain BALB/C)
C:Species: Mus musculus (house mouse)
C:Date: 19-Jan-1996 #sequence_revision 19-Jan-1996 #text_change 16-Aug-1996
C:Accession: A56446
R:Tang, P.M.; Foltz, L.A.; Mahoney, W.C.; Schueler, P.A.

J. Biol. Chem. 270, 7829-7835, 1995
A:Title: A high affinity digoxin-binding protein displayed on M13 is functionally identical
A:Reference number: A56446; MUID:95229583; PMID:7713873
A:Accession: A56446
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-268 <TAN>
A:Cross-references: GB:U20617
C:Keywords: heterotrimer; immunoglobulin

Query Match 29.6%; Score 694; DB 2; Length 268;
Best Local Similarity 55.2%; Pred. No. 2,1e-41;
Matches 133; Conservative 43; Mismatches 63; Indels 2; Gaps 2;

QY 1 QIQVVGSGELKPEETVYKISCKASGYPTNYGNMVAQAPGQGLKMMGINTSGESTF 60
DB 3 QYKQESGELVYKPSAVKLTCTTSGFNKDTYMWVWQKRPQGLMGRAPAGITKY 62
QY 61 ADPKGRPDSLETSAANTAYIQINNLSKEDMATYFCARMEVYHGVYVYWGQTTVYSSG 120
DB 63 DPKFQKATIAADTSSNTAYIQLSLTSEDTAVYVCASYLTR-YENYWGQTTVYSSG 121
QY 121 GGGSGGGGGGGSDIQLTQSHKFLSTVGDVRSITCKASQDYNAVAMVQKPEQSKL 180
DB 122 GGGSGGGGGGGSDIELTQSPALMSASLGEKVTWSCASSSV-NFIYVQKSPASPKL 180
QY 181 LIYSASRYTGVPSRFTSGSGGPDFTTISVQAEIDLAVYFCQGHFRTPFTFGSGTKLEI 240
DB 181 WYLYSHLPQVPAFNFSSGSGNSISLTISMEGDAATYCCQPTSSPFTFGSGTKLEI 240
QY 241 K 241
DB 241 K 241

RESULT 3

Single chain Fv antibody - mouse
C:Species: Mus musculus (house mouse)
C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 06-Jan-1995
C:Accession: S41374
R:Artsenko, O.; Weiler, E.W.; Muentz, K.; Conrad, U.
Submitted to the EMBL Data Library, January 1994
A:Description: Construction and functional characterization of a single chain Fv antibody
A:Reference number: S41374
A:Accession: S41374
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-249 <ART>
A:Cross-references: EMBL:Z29480

Query Match 28.4%; Score 665.5; DB 2; Length 249;
Best Local Similarity 52.0%; Pred. No. 1.9e-39;
Matches 128; Conservative 47; Mismatches 66; Indels 5; Gaps 1;

QY 1 QIQVVGSGELKPEETVYKISCKASGYPTNYGNMVAQAPGQGLKMMGINTSGESTF 60
DB 1 QVQLOQSGAEIVRPGASVKTCTASGFNFKDDYTHWVQKPEKLEWIAIAPASGVKY 60
QY 61 ADPKGRPDSLETSAANTAYIQINNLSKEDMATYFCARMEVYHGVYVYWGQTTVYSSG 120
DB 61 VRFQDQKATITADTSSNTAYQLSLTSEDTAVYVCARRDTLYSLGYNQSGSTVYSSR 120
QY 121 GGGSGGGGGGGSDIQLTQSHKFLSTVGDVRSITCKASQ-----DVNAVAMVQKPG 175
DB 121 GGGSGGGGGGGSDIELTQSPVAVVIPESVISISRSKSLISGDSYLTWFLQRP 180
QY 176 QSPKLLIYSASRYTGVPSRFTSGSGGPDFTTISVQAEIDLAVYFCQGHFRTPFTFGSG 235
DB 181 QSPDILLIYRMSNLASGVDRFSGSGSTFTLRISRVEADVGVYVCMQREYPLTFGAG 240
QY 236 TKLEIK 241
DB 236 TKLEIK 241

DB 241 TKLEIK 246

RESULT 4

p53 specific single-chain antibody pab421 - human
C:Species: Homo sapiens (man)
C:Date: 15-May-1997 #sequence_revision 15-May-1997 #text_change 18-Jul-1997
C:Accession: J05322
R:Janot, C.B.; Hynes, N.E.
Biochem. Biophys. Res. Commun. 230, 242-246, 1997
A:Title: Characterization of scFv-421, a single-chain antibody targeted to p53.
A:Reference number: J05322; MUID:97168950; PMID:9016757
A:Accession: J05322
A:Molecule type: mRNA
A:Residues: 1-233 <UAG>
A:Experimental source: hybridoma cell
C:Comment: This protein specifically binds the tumor suppressor protein p53. It restores

Query Match 26.1%; Score 611.5; DB 2; Length 233;
Best Local Similarity 51.0%; Pred. No. 1.1e-35;
Matches 122; Conservative 39; Mismatches 67; Indels 11; Gaps 3;

QY 6 QSGBELKPKPEETVYKISCKASGYPTNYGNMVAQAPGQGLKMMGINTSGESTFADPK 65
DB 2 ESGAEIVRSASVKTCTTSGFNINDYVMWVKKRPEQGLEWIGRIDPENGADMTRSSG 61
QY 66 GRPDSLETSAANTAYIQINNLSKEDMATYFCARMEVYHGVYVYWGQTTVYSSGGSGSG 125
DB 62 VKATMTADTSSNTAYQLSLTSEDTAVYCC-----NAGMDYWGQTTVYSSGGSGSG 115

QY 126 GGGSGGGGGSDIQLTQSHKFLSTVGDVRSITCKASQDY-----YNAVAMVQKPGSKL 181
DB 116 GRASGGGGSDIELTQSPALMSASLGEKVTWSCASVSTSGSYWMMQKRPQGL 175
QY 182 IYASASRYTGVPSRFTSGSGGPDFTTISVQAEIDLAVYFCQGHFRTPFTFGSGTKLEI 240
DB 176 IYLVSNLSEGVPAFNFSSGSGGDTFTLNIHVEEDDAATYCC-QHIRELRSRGTKLEI 233

RESULT 5

T-cell receptor α precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 08-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 05-Nov-1999
C:Accession: S54817
R:Nocentini, G.; Ronchetti, S.; Bartoli, A.; Testa, G.; d'Adamo, F.; Ricciardi, C.; Migli
Submitted to the EMBL Data Library, January 1995
A:Description: T cell receptor α : an alternatively spliced product of the T cell recep
A:Reference number: S54817
A:Accession: S54817
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-177 <MOC>
A:Cross-references: EMBL:X84237; NID:g809043; PIDN:CA59015.1; PID:g809046
C:Keywords: T-cell receptor

Query Match 25.8%; Score 605; DB 2; Length 177;
Best Local Similarity 99.2%; Pred. No. 2.1e-35;
Matches 117; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 305 LEDPKLCYLLDILTYGYIALYIRAKFSASATLANLOPNOLYNELNGRREEDV 364
DB 26 LDDPKLCYLLDILTYGYIALYIRAKFSASATLANLOPNOLYNELNGRREEDV 85
QY 365 LEKKQARDPEWGKQORRRNPQEGVYNALQDKMAEAYSIGTYGERRRGKHGDL 422
DB 86 LEKKQARDPEWGKQORRRNPQEGVYNALQDKMAEAYSIGTYGERRRGKHGDL 143

RESULT 6

A35900
T-cell receptor CD3 ϵ chain precursor - mouse

C:\Species: Mus musculus (house mouse)
C:\Date: 23-Oct-1990 #sequence revision 23-Oct-1990 #text_change 09-Jul-2004
C:\Accession: A35900; A60374; G46522
R:\Jin, Y.J.; Clayton, L.R.; Howard, F.D.; Koyasu, S.; Sieh, M.; Steinbrich, R.; Tarr, G.
Proc. Natl. Acad. Sci. U.S.A. 87, 3319-3323, 1990
A:\Title: Molecular cloning of the CD3eta subunit identifies a CD3zeta-related product in
F:\22-163/Product: T-cell receptor zeta chain #status predicted <MAT>
A:\Accession: A35900
A:\Status: preliminary
A:\Molecule type: mRNA
A:\Residues: 1-206 <JIN>
A:\Cross-references: UNIPROT:P29020; GB:M3158; NID:g192488; PIDN:AAA37398.1; PID:g309155
R:\Ohno, H.: Salto, T. Immunol. 2, 1117-1119, 1990
Int. Immunol. 2, 1117-1119, 1990
A:\Title: CD3zeta and eta chains are produced by alternative splicing from a common gene.
A:\Reference number: A60374; MUID:91190781; PMID:2150596
A:\Accession: A60374
A:\Status: not compared with conceptual translation
A:\Molecule type: DNA
A:\Residues: 144-206 <OHNS>
R:\Jensen, J.P.; Concalarelli, C.; Hou, D.; Relhanan, B.L.; Dean, M.; Weisman, A.M.
J. Immunol. 150, 122-130, 1993
A:\Title: T cell antigen receptor-eta subunit. Low levels of expression and limited cross
A:\Reference number: A46522; MUID:93107707; PMID:8417118
A:\Contents: annotation
C:\Comment: The functional significance of this alternatively spliced product of the CD3
ino acids, differs widely among various mammalian species in sequence, length, and even
C:\Keywords: alternative splicing; T-cell receptor; transmembrane protein

Query Match 25.8%; Score 605; DB 2; Length 206;
Best Local Similarity 99.2%; Pred. No. 2.6e-35;
Matches 117; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 305 LEDPRLCTLDGILFIYGVITLALYLRAKFSRSATAANLODPQLYNELNGRREEDV 364
|||
Db 26 LDPLRLCTLDGILFIYGVITLALYLRAKFSRSATAANLODPQLYNELNGRREEDV 85

OY 365 LEKKRPDPEWGGKQOORRRNPQEGVYNALQDKMAEAYSEIGTGERRRGKHDLGYQ 422
|||
Db 86 LEKKRPDPEWGGKQOORRRNPQEGVYNALQDKMAEAYSEIGTGERRRGKHDLGYQ 143

RESULT 7
A31768
T-cell receptor zeta chain precursor - human
C:\Species: Homo sapiens (man)
C:\Date: 07-Jun-1990 #sequence_revision 07-Jun-1990 #text_change 09-Jul-2004
R:\Weissman, A.M.; Hou, D.; Orloff, D.G.; Modi, W.S.; Seanez, H.; O'Brien, S.J.; Klausne
Proc. Natl. Acad. Sci. U.S.A. 85, 9709-9713, 1988
A:\Title: Molecular cloning and chromosomal localization of the human T-cell receptor zet
A:\Reference number: A31768; MUID:89071765; PMID:2974162
A:\Accession: A31768
A:\Molecule type: mRNA
A:\Residues: 1-163 <WEIS>
A:\Cross-references: UNIPROT:P20963; GB:J04132; NID:9623041; PIDN:AAA60394.1; PID:9623042
C:\Keywords: phosphoprotein; T-cell receptor; transmembrane protein
F:\22-163/Product: T-cell receptor zeta chain #status predicted <MAT>

Query Match 25.8%; Score 604.5; DB 2; Length 163;
Best Local Similarity 84.9%; Pred. No. 2.1e-35;
Matches 118; Conservative 6; Mismatches 14; Indels 1; Gaps 1;

OY 305 LEDPRLCTLDGILFIYGVITLALYLRAKFSRSATAANLODPQLYNELNGRREEDV 364
|||
Db 26 LDPLRLCTLDGILFIYGVITLALFLRKFSRSAPPYYQGQOLYNELNLGRREEDV 85

OY 365 LEKKRPDPEWGGKQOORRRNPQEGVYNALQDKMAEAYSEIGTGERRRGKHDLGYQGL 424
|||
Db 86 LDKKRRPDPEWGGK-PRIKNQDEGLYNELQDKMAEAYSEIGMGERRRGKHDLGYQL 144

OY 425 STATKVTDALHMQTLADR 443

Db 145 STATKDTYDALHMQALPPR 163

|||||

RESULT 8

A:Accession: A45089

T-cell receptor CD3 theta chain, alternate splice form - mouse

C:Species: Mus musculus (house mouse)

C:Date: 30-Apr-1993 #sequence_revision 18-Nov-1994 #ext_change 05-Nov-1999

C:Accession: A45089; 149587

R:Clayton, L.K.; Diener, A.C.; Lerner, A.; Tse, A.G.; Koyasu, S.; Reinherz, E.L.

J. Biol. Chem. 267, 26023-26030, 1992

A:Title: Differential regulation of T-cell receptor processing and surface expression aff

A:Reference number: A45089; MUID:93100325; PMID:1464613

A:Accession: A45089

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-188 <CLA>

A:Cross-references: GB:S51932; NID:g261998; PIDN:AA824559.1; PID:g261999

A:Experimental source: thymus

A:Note: Sequence extracted from NCBI backbone (NCBI:P:120865)

R:Lerner, A.; Diener, A.C.; Reinherz, E.L.; Clayton, L.K.

Eur. J. Immunol. 22, 2135-2140, 1992

A:Title: Human genomic sequences corresponding to murine CD3eta-related transcripts: lach

A:Reference number: 149587; MUID:92247411; PMID:1322304

A:Accession: 149587

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-188 <RES>

A:Cross-references: GB:L03353; NID:g192508; PIDN:AAA37401.1; PID:g192509

C:Keywords: T-cell receptor

Query Match 25.1%; Score 589.5; DB 2; Length 188;

Best Local Similarity 98.3%; Pred. No. 2.8e-34;

Matches 116; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Oy 305 LEDRLCTVLGGILFIYGVITLALYLAKFSRSAAETPAALQDPNOLYNELNLRREYDV 364

Db 26 LLDPLCTVLGGILFIYGVITLALYLAKFSRSAAETPAALQDPNOLYNELNLRREYDV 85

Oy 365 LEKKRAPDPKMGKQKQRRNPQEGVYNALQKQKMAEYSEIGTKGRRRKGGHGLYQ 422

Db 86 LEKKRAPDPKMGKQKRRNPQEGVYNALQKQKMAEYSEIGTKGRRRKGGHGLYQ 142

RESULT 9

JC4664

T-cell receptor zeta chain - bovine

C:Species: Bos primigenius taurus (cattle)

C:Date: 10-May-1996 #sequence_revision 19-Jul-1996 #ext_change 05-Nov-1999

C:Accession: JC4664

R:Hagens, G.; Galle, Y.; Glaser, I.; Davis, W.C.; Baldwin, C.L.; Clevers, H.; Dobbelaere

Gene 163, 165-171, 1996

A:Title: Cloning, sequencing and expression of the bovine CD3 epsilon and TCR-zeta chains

A:Reference number: JC4663; MUID:96194796; PMID:8647441

A:Accession: JC4664

A:Molecule type: mRNA

A:Residues: 1-166 <HAG>

A:Cross-references: GB:U55688; NID:g1263011; PIDN:AA048548.1; PID:g1263012

C:Comment: This protein plays a pivotal role in linking T-cell receptor-triggering to se

ymphokine receptor gene expression.

C:Genetics:

A:Gene: tcr-zeta

C:Keywords: GTP binding; signal transduction; T-cell receptor

F:123-146/Region: GDP/GTP-binding

Query Match 23.0%; Score 539; DB 2; Length 166;

Best Local Similarity 75.2%; Pred. No. 8.3e-31;

Matches 106; Conservative 11; Mismatches 22; Indels 2; Gaps 1;

Oy 305 LEDRLCTVLGGILFIYGVITLALYLAKFSRSAAETPAALQDPNOLYNELNLRREYDV 364

Db 26 LLDPLCTVLGGILFIYGVITLALYLAKFSRSAAETPAALQDPNOLYNELNLRREYDV 85

submitted to the EMBL Data Library, August 1993

A;Reference number: S37200

A;Reference number: S37200

A;Accession: S37204

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-118 <FIS>

A;Cross-references: EMBL:X74589; NID:g402599; PIDN:CAA52666.1; PID:g1333981

C;Superfamily: Immunoglobulin V region; Immunoglobulin homology

F;15-98/Domain: Immunoglobulin homology <IMM>

Query Match 22.3%; Score 522.5; DB 2; Length 118;

Best Local Similarity 80.7%; Pred. No. 7.8e-30;

Matches 96; Conservative 9; Mismatches 13; Indels 1; Gaps 1.

```
DQ      1 QIQLVSGPELKKPGETVTKISCAASGYPTNYGMNWKAPGGGLKMMGWNITSTBSTF   600  
       |||||                               :|||||  
DB      1 QIQLVSGPELKKPGETVKISCASGYTFIDYSNMWKPVGKLKMGNINTELTGAAY    600
```

Db 1 QIQLVQSGPELKKPGETVKISCKASGYTFIDYSMHVVKQVPCKGLKMMGWINTEGEAKY 600

```
Qy      61  ADDEKGRFDSLETSANTAYLQINLKSEDMATYFCARNEVHGVYPYWGOGTIVTVSS 113
      |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db      61  ADDFKGRFASLETSASTAYLQINLKNEEDTATYFCARNG-YSANFPYWGOGTLVTVSA 118
```

Db 61 ADDFKGRFAFSLETSASTAYLQINLNKEDTATYFCARWG-YSAWFPYWGQGLTVSA 118

RESULT 15

A53285

Ig heavy chain V and J regions, monoclonal antibody SCET.M8.1 - mouse (fragment)

C;Species: Mus musculus (house mouse)

C:\Date: 02-May-1994 #sequence_revision 18-Nov-1994 #text_change 20-Jun-2000

C;Accession: A53285

R; Sawada, J.; Mizusawa, S.; Terao, T.; Naito, M.; Kurosawa, Y.

Mol. Immunol. 28, 1063-1072, 1991

A; Title: Molecular characterization of monoclonal anti-steroid antibodies: primary struc

and their pH-reactivity profiles.

A;Reference number: A53285; MUID:92017897; PMID:1922102

A;Accession: A53285

A;Status: preliminary

A; molecule type: DNA; protein

A;Residues: 1-119 <SAW>

A; Cross-references: GB:D12736; NID:g220595; PIDN:BAA02228.1; PID:g220596

A:Note: sequence extracted from NCBI backbone (NCBIN:63271, NCBIP:63299)

C;Superfamily: immunoglobulin V region; immunoglobulin homology

C;Keywords: heterotetramer; immunoglobulin

F;15-98/Domain: immunoglobulin homology <IMM>

Query Match 22.1%; Score 519; DB 2; Length 119;

Best Local Similarity 82.4%; Pred. No. 1.4e-29;

Matches 98; Conservative 8; Mismatches 13; Indels 0; Gaps 0;

```

Oy 1 QIQLVQSGPELKKPGETVKISCKASGYPTNYGMNWKQAPGGGLKMGWINTSTGESETF 600
    |||||
Db 1 QIQLVQSGPELKKPGETVKISCKASGYPTNYGMNWKQAPGGGLKMGWINTSTGESEPT 600

```

Db 1 QIQLVSGPELKKPGETVKISCKASGYFTNYGMNVKQAPGKGLEWMDWINTYTGPTY 60

Qy	61	ADDEKGRFDESLETSANTAYLOINNLKSEDMATYFCAREVYHGVPYWGOTTVTWSS	119
		: : : : : :	
Db	61	ADDEKGRFASLETSASTAYLOINDLNKEDTATYFCARGNVRVYAMDYWGOSTVTWSS	119

Db 61 ADDEKGRFAFSLETSASTAYLQINDLKNEDTATYFCARGNVRVYAMDYWGQSTVSS 119

Search completed: April 20, 2005, 07:07:49
Job time : 18.1515 secs

Job time : 18.1515 sec

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 20, 2005, 10:13:45 ; Search time 4346.69 Seconds
(without alignments)
3879.380 Million cell updates/sec

Title: US-09-596-774-7
Perfect score: 2345
Sequence: 1 QIOLVOSGPELKKPGETVKI.....LSTATKDYALHMOTIAPR 443

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model DBV=xlh
-Q=/cg12_1/USPTO.spool/h/US09596774/runatc.20042005.075443.12521/app_query.fasta_1.1230
-DB=EST -QPM=fastap -SUFF1=apr20.tst -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=45
-UNITS-bits -START=1 -END=1 -MATRIX=blousum62 -TRANS=human40.cdi -LIST=45
-DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTPM=pic -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USRR=US09596774.@CGN 1.1_4591.@runatc.20042005.075443.12521 -NCPV=6 -ICPV=3
-NO_MMAP -LARGESUBERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1.*
2: gb_est2.*
3: gb_hc.*
4: gb_est3.*
5: gb_est4.*
6: gb_est5.*
7: gb_est6.*
8: gb_g981.*
9: gb_g982.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	746.5	31.8	671	AF240168	AF240168 Mus muscu
2	715	30.5	953	BY746048	BY746048 BY746048
3	715	30.5	1570	AK088087	AK088087 Mus muscu
4	713	30.4	577	CA577711	CA577711 K0713C05
5	704.5	30.0	1598	AK017904	AK017904 Mus muscu
6	697.5	29.7	492	AY403844	AY403844 Mus muscu
7	681.5	29.1	764	C0568334	C0568334 AGENCOURT
8	681.5	29.1	786	C0572391	C0572391 AGENCOURT
9	628	26.8	993	BI832527	BI832527 603082128

c	10	621	26.5	872	7	CK629396	CK629396 AMO-AA001
	11	620	26.4	618	7	CV029513	CV029513 8329 Full1
	12	620	26.4	772	4	BI838213	BI838213 603083131
	13	620	26.4	1038	5	BM921412	BM921412 AGENCOURT
	14	620	26.4	1113	1	AL554350	AL554350 AL554350
	15	620	26.4	1546	3	CR625818	CR625818 full-1eng
	16	620	26.4	1557	3	CR601423	CR601423 full-1eng
	17	617.5	26.3	547	5	B0562923	B0562923 H4078G04-
	18	615	26.2	1025	1	AL558350	AL558350 AL558350
	19	614	26.2	1025	1	AL558350	AL558350 AL558350
	20	607.5	25.9	899	7	CO648852	CO648852 ILLUMIGEN
	21	606.5	25.9	601	5	B0474958	B0474958 carabua4g
	22	604.5	25.8	492	9	AY403842	AY403842 Homo sapi
	23	604.5	25.8	575	5	BP367299	BP367299 BP367299
	24	604.5	25.8	581	5	BP368604	BP368604 BP368604
	25	604.5	25.8	712	1	AL557555	AL557555 AL557555
	26	604.5	25.8	875	5	EX464301	EX464301 BX464301
	27	604.5	25.8	970	5	B0051777	B0051777 AGENCOURT
	28	604.5	25.8	1095	1	AL532484	AL532484 AL532484
	29	604.5	25.8	1126	3	CR615998	CR615998 full-1eng
	30	602.5	25.7	1104	7	CK629846	CK629846 AM2-AA002
c	31	596.5	25.4	580	5	BP367951	BP367951 BP367951
	32	578.5	24.7	581	5	BP366659	BP366659 BP366659
	33	567.5	24.2	852	4	BI819337	BI819337 603035487
	34	567	24.2	672	7	CK633068	CK633068 AM3-AP001
	35	567	24.2	672	7	CK633068	CK633068 AM3-AP001
	36	565	24.1	675	5	BP163731	BP163731 BP163731
	37	565	24.1	707	5	BP167815	BP167815 BP167815
	38	565	24.1	727	5	BP164018	BP164018 BP164018
	39	565	24.1	823	5	BP167326	BP167326 BP167326
	40	565	24.1	835	5	BP163661	BP163661 BP163661
	41	565	24.1	841	5	BP161017	BP161017 BP161017
	42	565	24.1	841	5	BP164042	BP164042 BP164042
	43	565	24.1	845	5	BP434480	BP434480 BP434480
	44	565	24.1	885	5	BP161782	BP161782 BP161782
	45	564	24.1	695	5	BP166127	BP166127 BP166127

ALIGNMENTS

RESULT 1
AF240168
LOCUS AF240168 671 bp mRNA linear HTC 30-APR-2001
DEFINITION Mus musculus MRP5 mRNA, partial cds.
ACCESSION AF240168
VERSION AF240168.1 GI:13877288
KEYWORDS
SOURCE
ORGANISM Mus musculus (house mouse)
REFERENCE
AUTHORS Cui,D.X., Zeng,G.Y., Wang,F., Xu,J.R., Ren,D.O., Guo,Y.H., Tian,F.R., Yan,X.J., Hou,Y., Ren,D.O., Guo,Y.H.,
TITLE Mechanism of exogenous nucleic acids and their precursors improving the repair of intestinal epithelium after gamma-irradiation in mice
JOURNAL World J. Gastroenterol. 6 (5), 709-717 (2000)
PUBMED 11819679
REFERENCE
AUTHORS Cui,D., Zeng,G., Yan,X., Li,X. and Su,C.
TITLE Cloning of mouse genes related to repairing of intestinal epithelium of the irradiated mice by treatment with the intestinal RNA of mice of the same strain
JOURNAL Int. J. Radiat. Biol. Relat. Stud. Phys. Chem. Med. 19 (1), 71-80 (2001)
REFERENCE
AUTHORS Cui,D., Zeng,G., Yan,X., Wang,F., Tian,F., Ren,D., Zhao,T., Li,X. and Su,C.
TITLE Direct Submission
JOURNAL Submitted (02-MAR-2000) Molecular Cloning Laboratory, Institute of Genetic Diagnosis, Fourth Military Medical University, No. 17 Chang Le West Road, Xi'an 710032, China
FEATURES
Location/Qualifiers

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 A CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC Building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

FEATURES

location/Qualifiers
 1. 953
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="NOD"
 /db_xref="taxon:10090"
 /clone="E430003L03"
 /tissue_type="thymus"
 /cell_type="thymic cells"
 /clone_id="RIKEN full-length enriched, 2 days neonate thymus thymic cells (NOD)"

ORIGIN

Alignment Scores:
 Pred. No.: 9.36e-61 Length: 953
 Score: 715.00 Matches: 146
 Percent Similarity: 86.39% Conservative: 0
 Best Local Similarity: 86.39% Mismatches: 11
 Query Match: 30.49% Indels: 12
 DB: 6 Gaps: 2

US-09-596-774-7 (1-443) x BY746048 (1-953)

275 SerProValHisPrcThnGlyThrSerGlnProGlnArgProGluAapCyAspProArg 294
 94 TCGCTGCTGCTCTCCACCTGCGGCTCCAGAGCAG-----AGCAGCAGGA 138
 295 GlySerValIySgIyThnGlyLeuAapPheLeuGluAapProIyLeuCyAspTyLeuLeu 314
 139 GCTTGGCTCTG-----CTGATCCCAACTCTGCTACTTGTCTA 177
 315 AapGlyIleLeuPheIleTyGlyValIleIleThraIeLeuTyLeuAaGAlaIyPhe 334
 178 GATGGAATCCCTCTCATCTACGAGATCATCATCACGCCCTGTACTGAGAGCAAAATTC 237
 335 SerArgSerIaGluThraIaAlaAapLeuGlnAapProAaGlnIleuTyAsnGluLeu 354
 238 AGCAGAGTGAGAGACTGCTGCCAAGTCCAGAGCCCAACAGCTCTTACAAATGAGCTC 297
 355 AapLeuGlyArgArgGluGluTyAspValLeuGluIyIySvAspAlaArgAspProGlu 374
 298 AATTAAGGCGCAAGAGAGATATGACTCTTGAAGAAAGACGGGCTCGGATCCAGAG 357
 375 MetGlyIyIySgIyThnGlyLeuAaGArgAaPrcGlnGluGlyValTyAsnAlaLeuGln 394
 358 ATGGAGAGCAAAACGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 417
 395 LyAaPrcIyAaMetAlaGAlaIyTySerGlnIleGlyThryIySgIyGluArgArgGly 414
 418 AAGACAAAGATGGAG 477
 415 LySgIyIyIyAapGlyLeuTyGlnGlyLeuSerThraIaThryIyAapThyTyAspAla 434
 478 AAGGAGCAGATGCGCTTACCAAGGCTCTGAGCACTGCCACCAAGAGAGAGAGAGAG 537
 435 LeuHisMetGlnThryLeuAlaPrcArg 443
 538 CTGCATATGACAGACCTGGCCCTCTGCG 564

RESULT 3
 AK088087
 LOCUS
 DEFINITION
 1570 bp mRNA linear HTC 03-APR-2004
 Mus musculus 2 days neonate thymus thymic cells cDNA. RIKEN full-length enriched library, clone:E430003L03 product:CD3 antigen, zeta polypeptide, full insert sequence.

ACCESSION
 AK088087
 VERSION
 AK088087.1 GI:26104673
 KEYWORDS
 HTC; CAP trapper.
 SOURCE
 Mus musculus (house mouse)
 ORGANISM
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 AUTHORS
 Carninci, P. and Hayashizaki, Y.
 TITLE
 High-efficiency full-length cDNA cloning
 JOURNAL
 Meth. Enzymol. 303, 19-44 (1999)
 MEDLINE
 99279253
 PUBMED
 10349636

REFERENCE
 AUTHORS
 2
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 TITLE
 Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes
 JOURNAL
 Genome Res. 10 (10), 1617-1630 (2000)
 MEDLINE
 20499374
 PUBMED
 11042159

REFERENCE
 AUTHORS
 3
 Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kikunishi, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kaishiki, K., Fujiwara, S., Inoue, K., Togawa, Y., Irawa, M., Ohara, E., Matsumi, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.
 TITLE
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer
 JOURNAL
 Genome Res. 10 (11), 1757-1771 (2000)
 MEDLINE
 20530913
 PUBMED
 11076861

REFERENCE
 AUTHORS
 4
 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
 TITLE
 Functional annotation of a full-length mouse cDNA collection
 JOURNAL
 Nature 409, 685-690 (2001)
 MEDLINE
 11076861

REFERENCE
 AUTHORS
 5
 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
 TITLE
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 JOURNAL
 Nature 420, 563-573 (2002)
 MEDLINE
 11076861

REFERENCE
 AUTHORS
 6
 (Bases 1 to 1570)
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirose, T., Hirose, T., Horii, F., Imotani, K., Ishii, Y., Itoh, M., Kigawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Ohnato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Watanabe, M. and Hayashizaki, Y.
 TITLE
 Direct Submission
 JOURNAL
 Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.jp, URL: <http://genome.gsc.riken.jp/>, Tel:81-45-503-9222, Fax:81-45-503-9216)

COMMENT
 CDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics
Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome
Trust/MRC building Addenbrookes Hospital Cambridge) whose
assistance we gratefully acknowledge.
Please visit our web site for further details.
URL: <http://genome.gsc.riken.jp/>.
URL: <http://phantom.gsc.riken.jp/>.
Location/Qualifiers

FEATURES

source
1..1570
/organism="Mus musculus"
/mol_type="mRNA"
/strain="NOD"
/db_xref="FANTOM DB:E430003L03"
/db_xref="taxon:10090"
/clone="E430003L03"
/cell_type="thymic cells"
/tissue_type="thymus"
/clone_id="RIKEN full-length enriched mouse cDNA library"
/dev_stage="2 days neonate"
76..571
/note="CD3 antigen, zeta polypeptide (MGD|MG1:88334,
GB|NM_031162, evidence: BLASTN, 100%, match=536)
putative"
polya_signal
1549..1554
/note="putative"
polya_site
1570
/note="putative"

ORIGIN

Alignment Scores:
Pred. No.: 1..9e-60 Length: 1570
Score: 715.00 Matches: 146
Percent Similarity: 86.39% Conservative: 0
Best Local Similarity: 86.39% Mismatches: 11
Query Match: 30.49% Indels: 12
DB: 3 Gaps: 2

US-09-596-774-7 (1-443) x AK088087 (1-1570)

QY 275 SerProValHisProthnGlyThrSerGlnProGlnArgProGluAspCysArgProArg 294
DB 98 TCGGCTGATCCTCCACGTCGCGTTCCACAGAGCAG-----AGGCACAGA 142
QY 295 GlySerVallysglyThnrglyLeuaspPheleuGluaspProlysluGlyCysTyrleuLeu 314
DB 143 GCTTTGGCTCG-----CTGATCCCAAACTCTGCTACTCTCTA 181
QY 315 AspGlylleuPheleuPheleuVallelleThraIaleuTyrleuArgAlaIysPhe 334
DB 182 GATGATCCCTCTCATCTACGAGATCATCTACAGCCCTGTAAGAGCAAAATTC 241
QY 335 SerArgSerAlaGluThrAlaAlaAsnleuGlnaaspProanglnleuTyrAsnGluLeu 354
DB 242 AGCAGAGTGCAGAGACGTGCTGCCAAGCTGCAGAGCCCAACGCTCTCAATGAGCTC 301
QY 355 AsnleuGlyArgArgGluGluTyrAspValleuGluIysIysArgAlaAspProGlu 374
DB 302 AATCTAAGGCGCAAGAGAGATATGACCTCTGAGAAAGAGCGGCTCGGATCCAG 361
QY 375 MetGlylylsglnGlnArgArgAspProGlnGluGlyValTyrAsnAlaIeuGln 394
DB 362 ATGGAGGCAAAACGACAGAGAGAGAACCCCAAGAGCGGTATCAATGCACTGACAG 421
QY 395 LysAspIysMetAlaGluAlaTyrSerGluIleGlyThrIysGlyGluArgArgGly 414
DB 422 AAAGACAAGATGGAGAGCCTACAGATGAGTCGGCAAAAGCGAGAGCGGAGAGGC 481
QY 415 LysGlyIHisAspGlyLeuTyrGlnGlyLeuSerThraIthrIysAspThrTyrAspAla 434

DB 482 AAGGCGACATGAGCTTTACAGGCTTCAGACATGCCACAGACACCTATGATGCC 541
QY 435 LeuHisMetGlnThrleuAlaProArg 443
DB 542 CTGCATATGACAGACCTTGCCCTCGC 568

RESULT 4
LOCUS CA577711 577 bp mRNA linear EST 19-NOV-2002
DEFINITION K0713C05-5N NIA Mouse Hematopoietic Stem Cell (lin-/c-Kit-/Sca-1-)
IMAGE:30074812 5', mRNA sequence.
CA577711

ACCESSION CA577711.1 GI:25126102

VERSION EST.

KEYWORDS Mus musculus (house mouse)

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Other_ESTs: K0713C05-3
Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@igsn.gsc.nia.nih.gov
Plate: K0713 Row: C Column: 05
Seq primer: M13 Reverse
High quality sequence stop: 577
POLYA=No.

FEATURES

Location/Qualifiers
1..577

/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6NCr"
/db_xref="nlaEST:K0713C05-5N"
/db_xref="taxon:10090"
/clone="NIA:K0713C05 IMAGE:30074812"
/tissue_type="Hematopoietic Stem Cell"
/lin-/c-Kit-/Sca-1-)"
/dev_stage="Age approx. 10 weeks old"
/lab_host="DH10B"
/clone_id="NIA Mouse Hematopoietic Stem Cell"
/note="Vector: pSPORT (Invitrogen); Site 1: SalI; Site 2:
NotI; Mouse cDNA project by the Laboratory of Genetics,
National Institute on Aging (NIA), Intramural Research
Program, NIH (<http://19sun.gsc.nia.nih.gov/cDNA>). This is
a long-transcript enriched cDNA library (Ref. Genome Res.
11: 1553-1558 (2001). [PMID: 11544191]). Total RNAs were
obtained from Drs. Dennis Taub, Dan Longo (National
Institute on Aging, USA), Jonathan Keller (National Cancer
Institute, USA). Double-stranded cDNAs were synthesized
with an Oligo(dT) primer (Invitrogen:
5'-pACTAGTCTTAGATCGGAGCGGCCCTTTT-3') from
0.9 ug of total RNA, treated with T4 DNA polymerase, and
purified by ethanol-precipitation. The cDNAs were ligated
to lone-linker L1-SalI, purified by phenol/chloroform, and
separated from free linkers by Centricon 100. Then, the
cDNAs were amplified by long-range high fidelity PCR using
Ex Tag polymerase (Takara) with a primer SalI-S. The
products were purified by phenol/chloroform and Centricon
100. The cDNAs were digested with SalI and NotI enzymes
and cloned into SalI/NotI site of pSPORT plasmid vector.
The DH10B E. coli host was transformed with the ligation
mixture by the standard chemical method. The average
insert size is about 2.1 Kb. The library was constructed
by Yulan Piao (NIA)."

ORIGIN

Alignment Scores:

Pred. No.:	7,3e-61	Length:	577
Score:	713.00	Matches:	138
Percent Similarity:	99.28%	Conservative:	0
Best Local Similarity:	99.28%	Mismatches:	1
Query Match:	30.41%	Indels:	0
DB:	6	Gaps:	0

US-09-596-774-7 (1-443) x CM577711 (1-577)

Qy 305 LeuGIuApProLyLeuCyEtyrLeuLpGlyIleuPheIleTyrgIyValIle 324

Db 133 CTGCGATGCCCAAACTCTGCTACTGTGATGTAATCCTTCTATCTAGGAGTCAATC 192

Qy 325 ILeThralaLeuTyrlLeuAlaArgAlaYpSerAArgSerAlaGIuThAlAlaAenLeu 344

Db 193 ATCAAGCCCTGTACTCTAGAGCAAAATTCAGCAGAGATGAGACACTGCTGCCAACTG 252

Qy 345 GlnApPProAenGlnLeuTyraSngIuLeuAenGlnIyArGArGluGIuTyraSyaI 364

Db 253 CAGAGCCCAACCACTTACATATAGCTCAATCTTAGGCGAAGAGAAATATGACCTC 312

Qy 365 LeuGIuLyAlaArgAlaArgAapProGIuMetGIyGIySgInGlnArGArGhArgAn 384

Db 313 TTGGAGAAAGAAAGCGGCTCGGATCCAGATCGGAGCAAAAGCAGAGAGAGAGAAC 372

Qy 385 ProGIuGIuGIyValIyTyraSnaIleuGIuLyAapLyMetAlaGIuAlaTySerGIu 404

Db 373 CCCAGAGAAAGCGATATCAATGCACTGCAAGAAAGCAATGCGAGAGCTTACAGTGAG 432

Qy 405 ILeGIyThrylSgIyGIuArGArGArGIyLySgIyHSApGIyLeuTyGIuGIyLeu 424

Db 433 ATCGGCACAAAGCGAGAGCGAGAGAGCAAGGCGCGATGCGCTTACAGGAGTCTC 492

Qy 425 SerThralThrylAapThryTyraSnaIleuHismetGIuThLeuAlaProArg 443

Db 493 AGCACTGCCCAAGAGACACTATGATGCTCTGCTATGCAAGCCCTGCGCTTCCG 549

RESULT 5

AKO17904 1598 bp mRNA linear HTC 03-APR-2004

LOCUS AKO17904

DEFINITION Mus musculus adult male thymus cDNA, RIKEN full-length enriched library, clone:5830407D18 product:CD3 antigen, zeta polypeptide, full insert sequence.

ACCESSION AKO17904

VERSION AKO17904.1 GI:12857392

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE 1 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 2 Carninci, P. and Hayashizaki, Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999)

JOURNAL MEDLINE 99279253

PUBMED 10349636

REFERENCE 3 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)

JOURNAL MEDLINE 20493374

PUBMED 11042159

REFERENCE 4 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watabiki, M.,

TITLE

JOURNAL MEDLINE 20530913

PUBMED 11076861

REFERENCE 4 The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)

5 The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)

6 (bases 1 to 1598)

Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Arai, A., Aono, H., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kaenaka, T., Kato, H., Kawai, J., Kojima, Y., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ono, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Tanahashi, F., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

TITLE

JOURNAL MEDLINE 20530913

PUBMED 11076861

REFERENCE 4 Submitted (10-JUL-2000) Yoshihide Hayashizaki. The Institute of Physical and Chemical Research (RIKEN) Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

cdna library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTVN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5'-GAGAGAGAGATTCGATTAATTAATTCCTCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3' end: BamHI. Host: DH10B.

FEATURES

SOURCE

1.1598

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="PANTOM DB:5830407D18"

/db_xref="taxon:10090"

/clone="5830407D18"

/sex="male"

/tissue_type="thymus"

/clone_id="RIKEN full-length enriched mouse cDNA library"

/dev_stage="adult"

112..606

/note="unnamed protein product: CD3 antigen, zeta polypeptide (MGDI:88334, GB|NM_031162, evidence: BLASTN, 100%, match=536)

putative"

CDS

	/codon_start=1	
	/protein_id=BAB30997.1"	
	//db_xref=GI:12857393."	
	/translation="MKMRYSLVACLILHVRFGAEAKQSFLDLPKCYLLDGILEIFYGV ITITALYLAKEFSRSSETALQDPNQLNEINLGRREYDYLEKRADPEMGKKOQR RNPGEGYNNALOXRMAEAYSIEIGTKERRRRKGHDLYGLSTATKDTCDAIHMQC LAPR"	
polyA_signal	.1589	
polyA_site	/note="putative" .1598 /note="putative"	
ORIGIN		
Alignment Scores:		
Pred. No.: 2,196-59	Length: 1598	
Score: 704.50	Matches: 153	
Percent Similarity: 78.0%	Conservative: 3	
Best Local Similarity: 76.50%	Mismatches: 20	
Query Match: 30.04%	Indels: 24	
DB: 3 Gaps: 6		
US-09-596-774-7 (1-443) x AK017904 (1-1598)		
Oy ThrllyPProvalLeuaArThrProSer-----ProValHis 278		
Dd 9 ACAAAGCCACGA---GAGACTTCAATCACGCCTCCTTTCTCATCTCGGCCAGATA 65		
Oy 279 ProThrGlYltnSer-----GlnPrroInatrgProgluaap----- 290		
Dd 66 GTGGCTGTGCTTGCTTGCTCGGTGATACCATTCCCAGAGAACAAGATGAAGTGAAGT 125		
Oy 291 ---CySarq-----ProArgIlgySeVallysgllrghylleuaspp--phe-- 304		
Dd 126 GTCTGTTCTCGCTGCATCTTCACAGTCGGGTTCCAGAGAACAGAGGACAGACTTTGG 185		
Oy 305 -lenGuiaSpProlylsieucysrrtleuleuaspgilylleuehellerytglyyalii 324		
Dd 186 TCTGCTGGATCCAAACTCTCTACTCTGTAAGTAGAATCTCTTCAATTAACGAGATCAT 245		
Oy 324 eilethrAlaleutyrtleuarGalalyPhseSargseralagluThrAlalaenle 344		
Dd 246 CATCACAGCCTGTATCTGAAGAGCAAAATTCACAGAGATGCACAGACTGCCCCAACCT 305		
Oy 344 uGlnAspProasnglinleutyrasnGuluLeuanleuglyARgARGluTurYaarPa 364		
Dd 306 GCAGAGACCCCAACAGACTCTCAATGAGCTCAACTGAGGGGGAAGAGAGATATGACGT 365		
Oy 364 lLeugJullysyarsgalaaRGASPPROglmEtGlylglySGlnglnarGARgarGaS 384		
Dd 366 CTTGGAGAGAAGAGCGGCTCGGATCCAAGATGGAGGCAAAACACAGAGAGAGAGAAA 425		
Oy 384 nProGIngUlylaTyraenalAleuginLyaspLysMerAlaglualaryTsErGi 404		
Dd 426 CCCCCAGAGAGGCGTATATCAATGACTGCGAGAAAGACAMAAGTGCAGAGCCTTAcaGTga 485		
Oy 404 ulleGjThrllyselyguinarGarGaygilvylgcyhiiasprglyeurytgningyle 424		
Dd 486 GATGGGCACAAAAAGCAGAGGCGGAGAGGCAAGAGGCACAGATGGCTTTTAACAGGGTCT 545		
Oy 424 usErrThralathrlysaSPThrtYzaEPAlaleuhimelGInthrleualaProarg 443		
Dd 546 CAGACATGCCACCAAGAGCACCTGTGATGCTGCTTCATATGCAGACACCTGGCCCTTCGC 603		
RESULT 6		
Locus AY403844		
DEFINITION	AY403844 .492 bp DNA linear GSS 15-DEC-2003	
LOCUS	Mus musculus CD3z gene, VIRUAL TRANSCRIPT, partial sequence,	
ACCESSION	AY403844	
VERSION	AY403844.1 GI:39759827	
KEYWORDS	GSS.	
SOURCE	Mus musculus (house mouse)	
ORGANISM	Mus musculus	

REFERENCE		eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Scurionathii; Muridae; Murinae; Mus.
AUTHORS		1 (bases 1 to 492)
TITLE		Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejaritwal,A., Todd,M.A., Tanenbaum,D.M., Civeello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
JOURNAL		Inferred nonneutral evolution from human-chimp-mouse orthologous gene trios
PUBMED		Science 302 (5652), 1960-1963 (2003)
REFERENCE		14671302
AUTHORS		2 (bases 1 to 492)
TITLE		Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejaritwal,A., Todd,M.A., Tanenbaum,D.M., Civeello,D.R., Lu,F., Murphy,B., Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J., Adams,M.D. and Cargill,M.
JOURNAL		Direct Substition Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT		This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES		Location/Qualifiers
SOURCE		1..492
gene		/organism="Mus musculus"
		/mol_type="genomic DNA"
		/db_xref="taxon:10090"
		<1...>492
		/gene="CD32"
		/locus_tag="HCM1692"
ORIGIN		
Alignment Scores:		
Pred. No.:	2, 08e-59	Length: 492
Score:	697.50	Matches: 137
Percent Similarity:	98.56%	Conservative: 0
Best Local Similarity:	98.56%	Mismatches: 1
Query Match:	29.74%	Indels: 1
DB:	9	Gaps: 1
US-09-596-774-7 (1-443) x AY403844 (1-492)		
OY	305	LeuGluAspProIyLeuCyStyTleuLeuAspGlylleuPheileTygIyValille 324
Dd	76	CtGGTGGAATCCCAACTGTCTACCTGGTAGATGAACTCCTTCATCTACCGAGCATC 135
OY	325	IleThrAlaLeuTytleuArgAlaIysPheSerzrgeralagIunhrIalaAseu 344
Dd	136	ATCACAGCCCTTACTTACAGCAAAAATTGCAGAGATCAAGATCTGCTGCCAACCTG 195
OY	345	GlnAspProAngInLeuTyFasngIleuAseuIyArgrgsIngIntyrAspVal 364
Dd	196	CAGGACCCCACACAGCTCTTAACAATGACTCAATCTAGGCGAAGAAGAAATATGACGTC 255
OY	365	LeuGluLysArgAlaIarAspProGluMerGlyLyIsIngInIargrArgrasn 384
Dd	256	TTGGAGAAAGAGCGGGCTCGGGAATCCCAAAATGGAGGCATA--CAGAGGAGGGAAC 312
OY	385	ProGIngluGIyValTyFasnaIleuGIAltyAspIysMetIagIuaIatyrSeGlu 404
Dd	313	CCCCAGAAAGGGATTAACAATGACTGCTCAAAAAGACAAGATGGCAAAACCTTAAGTAG 372
OY	405	IleGlyThrIySGlyGluArgrArgrGLyLyGlyVHisAspIyleuTyGInglyLeu 424
Dd	373	ATCGGACAAAAAGGAGAGCGGAGAGGCGAAGGCGACGATGGCCCTTTACAGAGGCTTC 432
OY	425	SerThrIatThrIyAspThrTyFaspaIleuHiMetGIthrIeuAlaProArg 443
Dd	433	AGCACTGCCACCAAGACACCTATGATGCCCTGCATATGCAACCCTGACCCTGCC 489
RESULT 7		
LOCUS	COSE68334	764 bp mRNA linear EST 19-JUL-2004
DEFINITION	AGENSCOURT 286226036 NIH MG 250 Rattus norvegicus cDNA clone	

IMAGE:7384006 5', mRNA sequence.
 ACCESSION COS68334
 VERSION COS68334.1 GI:50380963
 KEYWORDS EST.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 764)
 NIH-MGC http://mgc.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics / NIH
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: c9apbs-remail.nih.gov
 Tissue Procurement: Drs. Josef Lazar & Howard Jacob, Medical
 College of Wisconsin
 CDNA Library Preparation: Open Biosystems
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLM15548 row: d column: 20
 High quality sequence stop: 746.
 Location/Qualifiers
 1..764
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /clone="IMAGE:7384006"
 /lab_host="DH10B Tona"
 /note="Organ: thymus; Vector: pExpress-1; Site: 1: EcorV;
 Site 2: NotI; RNA obtained from testis tissue of 8 wk old
 animal. Tissues were snap-frozen and kept at -80C before
 RNA extraction and purification (TRI-reagent method). CDNA
 was primed using oligo-dT primer:
 5'-pGACTAGTCTAGATCGGAGCGGCCGC(T)25-3' and cloned into
 the EcorV/NotI sites of pExpress-1. Size-selection >1.25kb
 resulted in an average insert size of 1.9 kb. This is a
 primary library (normalized library is NIH_MGC_251) and
 was constructed by Open Biosystems. Note: this is a
 NIH_MGC library"
 ORIGIN
 Alignment Scores: Length: 764
 Pred. No.: 681.50 Matches: 140
 Score: 81.56% Conservative: 6
 Percent Similarity: 78.21% Mismatches: 21
 Best Local Similarity: 29.06% Indels: 12
 Query Match: 7 Gaps: 3
 DB: 7
 US-09-596-774-7 (1-443) x COS68334 (1-764)
 Oy 274 ProserProvalHisPro-----ThrgLYThrsErlnPro 285
 Db 28 CCGCTCTGGGACACCTCCAGGTTGCAGAAATGAAGTGAGCGCATCGCTCGC 87
 Oy 286 GlnAArgProGluAArgCyAArgProAArgGlySerValLySGlyThrgLYleu-AapPheLe 305
 Db 88 CTGCATCTCTCAAGTCAAGTCCCGAGAGC-----AGAGCGACAGACCTTGGCTGCT 141
 Oy 305 uGlnAserProLyseuCyetyrLeuLeuAAspGlyIleleuPheIleTyrgLYValIleI 325
 Db 142 G---GATCCCAAACTCTCTATATGCTATGATGAAATCTCTTCATCTACGAGATCATGCT 198
 Oy 325 eThralaleuTyrlleuAgaIalalyPheSerAArgSerIaGluThralaIaAAsmLeuG1 345

Db 199 CACGGCCCTGTACTGAGACGAAATTCAGAGGAGTGCAGATCTGCTTACTTCA 258
 Oy 345 nAAsProAAsnGluInleuTyraAsnGluLeuAAsmLeuGlyAArgGluGluTyraAspValle 365
 Db 259 GGACCCCAACAGCTCTTATACGAGCTCAACTTACCTTACCTTACCTTACCTTACCTTACCTT 318
 Oy 365 uGluLyLeuAArgAlaAArgAAspProGluMetGlyGlyLyseGlnGlnAArgAArgAAsnPr 385
 Db 319 GCACAGAGAGCGGCTCGGAGTCCGAGATGCGGGGCAAAACAGACGAGAGAGAGAAACCC 378
 Oy 385 oGlnGlnGlyValTyraAsnAlaLeuGlnLyAAspLyMetAlaGluIleTyrsErGlnI 405
 Db 379 CCAGAAAGCGCTGTACATGATGATGCGAAAGACAAATGCGACAGGCTTACACTGAGAT 438
 Oy 405 eGlyThrgLYleuGlyAArgAArgGlyLyseGlyLyseAAspGlyLeuTyrgLYGlnLyLeu 425
 Db 439 TGGCATGAAGGCGAGAGCGGAGAGGCGGAGGCGGAGGCGGCTTTACCGAGGTTCCAG 498
 Oy 425 rThralaThrgLYAAspThrgLYAAspAlaLeuHisMetGlnThrgLYleuAlaProArg 443
 Db 499 CACTGCCACAGACGACCTATGACGCCCTGTCATATGACAGCCCTGCCCTGCC 553
 RESULT 8
 ACCESSION COS72391 786 bp mRNA linear EST 19-JUL-2004
 DEFINITION AGENCOURT 28621196 NIH_MGC_250 Rattus norvegicus CDNA clone
 LOCUS IMAGE:7381111 5', mRNA sequence.
 VERSION COS72391.1 GI:50385020
 KEYWORDS EST.
 SOURCE Rattus norvegicus (Norway rat)
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 786)
 NIH-MGC http://mgc.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics / NIH
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: c9apbs-remail.nih.gov
 Tissue Procurement: Drs. Josef Lazar & Howard Jacob, Medical
 College of Wisconsin
 CDNA Library Preparation: Open Biosystems
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLM15540 row: 1 column: 05
 High quality sequence stop: 701.
 Location/Qualifiers
 1..786
 /organism="Rattus norvegicus"
 /mol_type="mRNA"
 /db_xref="taxon:10116"
 /clone="IMAGE:7381111"
 /lab_host="DH10B Tona"
 /note="Organ: thymus; Vector: pExpress-1; Site: 1: EcorV;
 Site 2: NotI; RNA obtained from testis tissue of 8 wk old
 animal. Tissues were snap-frozen and kept at -80C before
 RNA extraction and purification (TRI-reagent method). CDNA
 was primed using oligo-dT primer:
 5'-pGACTAGTCTAGATCGGAGCGGCCGC(T)25-3' and cloned into
 the EcorV/NotI sites of pExpress-1. Size-selection >1.25kb
 resulted in an average insert size of 1.9 kb. This is a
 primary library (normalized library is NIH_MGC_251) and
 was constructed by Open Biosystems. Note: this is a
 NIH_MGC library"

ORIGIN

Alignment Scores:

Pred. No.:	1.62e-57	Length:	786
Score:	681.50	Matches:	160
Percent Similarity:	81.56%	Conservative:	4
Best Local Similarity:	78.21%	Mismatches:	21
Query Match:	29.06%	Indels:	12
DB:	7	Gaps:	3

US-09-596-774-7 (1-443) X C0572391 (1-786)

VERSION CK629396.1 GI:45753871
 KEYWORDS EST.
 SOURCE Apis mellifera (honey bee)
 ORGANISM Apis mellifera
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea; Apidae; Apis.
 1 (bases 1 to 872)
 Nunes, F.M.F., Valente, V., Sousa, J.F., Cunha, M.A.V., Pinheiro, D.G., Maia, R.M., Araújo, D.D., Costa, M.C.R., Martins, W.K., Carvalho, A.F., Mones, N., Nascimento, A.M., Peixoto, P.M.V., Silva, M.F.R., Ramos, R.G.P., Reis, L.F.L., Dias Neto, E., Souza, S.J., Simpson, A.J.G., Zago, M.A., Soares, A.E.E., Bitondi, M.G., Espinola, E.M., Espindola, F.S., Peco-Larson, M.L., Simoes, Z.L.P., Hartfelder, K. and Silva, M.A., Jr.
 TITLE Open reading frame ESTs - an efficient strategy for analysis of the honey bee transcriptome
 JOURNAL Unpublished (2004)
 COMMENT Contact: Silva Jr, W. A. Molecular Genetic and Bioinformatics Laboratory Department of Genetics, FMRP/USP, FUNDHERP Rua Tenente Celso Roxo, 2501, CEP 14051-140, Brazil Tel: +55 16 39639300 Fax: +55 16 39639309 Email: wilsonj@usp.br
 This sequence was derived from the FAPESP Genome Program High quality sequence start: 66
 High quality sequence stop: 625.
 Location/Qualifiers
 1..872
 /organism="Apis mellifera"
 /mol_type="mRNA"
 /strain="Africanized"
 /db_xref="taxon:7460"
 /sex="female, worker"
 /dev_stage="adult"
 /clone_id="AA0013"
 /note="Organ: whole body"

ORIGIN
 Alignment Scores:
 Pred. No.: 2,15e-51 Length: 872
 Score: 621.00 Matches: 122
 Percent Similarity: 71.43% Conservative: 38
 Best Local Similarity: 54.46% Mismatches: 63
 Query Match: 26.48% Indels: 4
 DB: 7 Gaps: 1

US-09-596-774-7 (1-443) x CK629396 (1-872)

QY 18 VallylleserCylylaserGlylPropherThaenlyrglymetAantTtVal 37
 Db 775 ATGAAATGCTCTCAAGCT-TCTGC-TACATCTTCACAAAGTATGATATGACTGGGTG 718
 QY 38 LysGlnAlaPrgGlyGlnGlyLeuLysTrpMetGlyTrpIleAenThSerThGlyGlu 57
 Db 717 AGGCAGACGCTGAACAGGAGCTTGAGTGGATGTTT-CCGAGAGAGGAGAGT 659
 QY 58 SerThrPheAlaAapPheLysGlyArpPheAapPheSerLeuGluThrSerAlaAa 77
 Db 658 ACTGAATACATGAGAGAGTCAAGGGAGGCCACACATGATGACCAAGTCCCTCCAGC 599
 QY 78 ThrAlaTrpLeuGlnIleAenAenLeuLysSerGlyAapMetAlaTrpPheCyAla 97
 Db 598 ACACCTATATGAGAGCTCAGCTGACATCTAGAGCTGCTGCTATTTCTGTGCT 539
 QY 98 ArgTrpGluValTyrHleGlyTyrValProlTyrTrpGlyGlnGlyThrThValThVal 117
 Db 538 AGAGGAGACTACTATAGGCGCTACTTTGACTTTGGGACCAAGGACCAAGGACCTC 479
 QY 118 SerSerGlyGlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlySerAapIleGln 137
 Db 478 TCCTCATGTGAGAGCGGTTGAGCGGAGGTGGCTCTGGCGGTGGCGGATCTGACATTGAG 419

QY 138 LeuThrGlnSerHisLysPheLeuSerThSerValGlyAapArgValSerIleThCys 157
 Db 418 CTCACCAAGTCTCCAGCAATATCTGCATCTCCAGGGAGAGGTCACCATGACCTGC 359
 QY 158 LysAlaSerGlnAapValTyrAenAlaAlaTrpTyrGlnGlnLysPrgGlyGlnSer 177
 Db 358 AGTCCAGCTCAAGTATA---CGTTACATATATGATGACCAAGAACCCGAGATCCCTCC 302
 QY 178 ProlYsLeuIleuLetyrSerAlaSerArgTyrThrGlyValProSerArgPheThr 197
 Db 301 CCCAGACTCCGATTATGACACATCCAGCTGCTCTGAGATCCCTTTTCGCTTCAGT 242
 QY 198 GlySerGlySerGlyProAapPheThrPheThrIleSerSerValGlnAlaGluAapLeu 217
 Db 241 GGAGTGGGTCTGGAGACTTATCTCTACATCAACGAATGAGAGGCTGAGAGTCT 182
 QY 218 AlaValTyrPheCyGlnGlnHisPheArgThrProPheThrPheGlySerGlyThrLys 237
 Db 181 GCCACTTATTACTGCCAGAGTGGAGTGTATCTGTACACGTTTCGAGAGGGGACCAAG 122
 QY 238 LeuGluIleLys 241
 Db 121 CTGAGGCTGAAA 110

RESULT 11
 CVO29513 618 bp mRNA linear EST 20-AUG-2004
 LOCUS sapiens cDNA 5' similar to BC025703, mRNA sequence.
 DEFINITION
 ACCESSION CVO29513
 VERSION CVO29513.1 GI:51487684
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 618)
 Rual, J.F., Hirozane-Kishikawa, T., Hao, T., Bertin, N., Li, S., Driscoll, A., Li, N., Rosenberg, J., Lamesch, P., Vidalain, P.O., Clingingsmith, T.R., Hartley, J.L., Espósito, D., Cheo, D., Moore, T., Simmons, B., Sequerra, R., Bosak, S., Doucette-Stamm, L., Le Peuch, C., Vandenhaute, J., Quick, M.E., Albaladejo, J.S., Hill, D.E. and Vidal, M.
 Human ORFome Version 1.1: a Platform for Reverse Proteomics
 Genome Res. (2004) In press
 CONTACT: Vidal M
 Marc Vidal Laboratory
 Dana Farber Cancer Institute
 1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA
 Tel: 617 632 5180
 Fax: 617 632 5739
 Email: Marc.Vidal@dfci.harvard.edu
 ORF Sequence Tag (OST) of Gateway Entry construct. Each cloned ORF results from a PCR reaction using an MGC full-length cDNA as template DNA and ORF specific primers
 PCR primers
 FORWARD: ATGACGTGAAGGCGCTTT
 BACKWARD: TACGACAGGGGAGGAGCTTG
 Insert Length: 618 Std Error: 29.00
 Plate: 11027 row: 09 column: H
 Seq primer: ACTGGCGCTGTTTACACAGCTGCTGACTGAGGAGAAAC
 High quality sequence start: 100
 High quality sequence stop: 617
 POLYANNO.

FEATURES
 source
 1..618
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /tissue="mixed"
 /clone_id="Full Length cDNA from the Mammalian Gene Collection"
 /note="Vector: mixed; The ORFs were PCR amplified from the

Alignment Scores:

1.66e-51	Length:	618
650.00	Matches:	126
79.64%	Conservative:	7
75.45%	Mismatches:	21
26.44%	Indels:	13
7	Gaps:	2

x CV029513 (1-618)

GlycylthreSerGlnProGlnArgProGlnAspCysArgProArgGlySerVal 29
 GGCACAGTTGCCGATTACAGA----- 59

GlyLeu-AspPheLeuGluAspProlyLeuGlyTyrLeuLeuAspGlyI 317
 |||||
 SAGCTTTGGGCTGCTG---GATCCAACTCTGCTACTGCTGATGGAAT 113

173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525
 526
 527
 528
 529
 530
 531
 532
 533
 534
 535
 536
 537
 538
 539
 540
 541
 542
 543
 544
 545
 546
 547
 548
 549
 550
 551
 552
 553
 554
 555
 556
 557
 558
 559
 560
 561
 562
 563
 564
 565
 566
 567
 568
 569
 570
 571
 572
 573
 574
 575
 576
 577
 578
 579
 580
 581
 582
 583
 584
 585
 586
 587
 588
 589
 590
 591
 592
 593
 594
 595
 596
 597
 598
 599
 600
 601
 602
 603
 604
 605
 606
 607
 608
 609
 610
 611
 612
 613
 614
 615
 616
 617
 618
 619
 620
 621
 622
 623
 624
 625
 626
 627
 628
 629
 630
 631
 632
 633
 634
 635
 636
 637
 638
 639
 640
 641
 642
 643
 644
 645
 646
 647
 648
 649
 650
 651
 652
 653
 654
 655
 656
 657
 658
 659
 660
 661
 662
 663
 664
 665
 666
 667
 668
 669
 670
 671
 672
 673
 674
 675
 676
 677
 678
 679
 680
 681
 682
 683
 684

caaaataaashnleugltnasprfioasnglnleutyAsnglnleuaenleugl 357
||| |||
ccccccgctaccagcaggcgccagaaacagactctatAACGAGCTCAATCTTAGG 233

[illegible][illegible]

GGCTACAGTAGATTGGGATGTAAGCGAGCCCGGAGCGCCAAAGGGCA 413

TTACACGGGTCTAGTACAGCCACCAAGGACACCTACGACGCCCTTCACAT 473

||||| 492
CCCCCTGC 492

772 bp mRNA linear EST 04-OCT-200

GI:15949763

Mollusca; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinoptera; Clupeiformes; Clupeidae; Sardinops
melastoxus; Sardinops melastoxus

to 772)
://mgc.nci.nih.gov/.
stitutes of Health, Mammalian Gene Collection (MGC)
(1999)
ert Strausberg, Ph.D.

Requirement: Life Technologies, Inc

Source

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
Plate: LLM11559 row: m column: 07
High quality sequence stop: 767.

1. .772

1. .772

```
/organism="Homo sapiens"
/mol_type="mpna"
```

/db_xref="taxon:9606"

```
/clone="IMAGE:5222574
/10b host="DVT00"
```

```
/lib/libc="DHLUB"
/clone_lib="NIH MGC 120"
```

/notes="Organ: pooled pancreas and spleen; Vector:

pcmv-spork16; site_1: NotI; site_2: EcoRV (destroyed); RNA source anonymous pool of arleen and nancton from 38 w/o

male. Library is oligo-dT primed and directionally cloned

(EcorV site is destroyed upon cloning). Average insert

normalized and enriched for full-length clones and was size 1.3 kb, insert size range 1-2.5 kb. Library is

constructed by C. Gruber (Invitrogen). Research Genetics

Tracking code 025. Note: this is a NIH_MGC Library.

Alignment Scores:

Pred. No.:	2 28e-51	Length:	772
Score:	620.00	Matches:	126
Percent Similarity:	79.64%	Conservative:	7
Best Local Similarity:	75.45%	Mismatches:	21
Query Match:	26.44%	Indels:	13
DB:	4	Gaps:	2

US-09-596-774-7 (1-443) X BI838213 (1-772

QY 278 HSPROTIN[GLYTHRserGIINProGI]NARProGIuaspCysargProargIySerVal 297
 ||||| |||||
 Db 90 CATCCTGCAGGCAcAGTTCcCGATTACAGa----- 119

Dy 298 LysGIYThrGlyLeu-AspPheLeuGlusppProLysLeuCystyrLeuLeuAaspGLyl] 317
 |||||
Dd 120 ---GGCACAGACCTTGGCGCTGCg---GATCCCAACTCtGTCTACTGCTGATGATGAAAT 173

Qy 317 leuPheIleTyGlyValIleIleThraIeaurTyIleuAargAlaIysPheSerArgSe 337
 |||||
 Db 174 CCTCTTCACTATGATGGTGCATTCTCACTG:CTTGTTCCTAGAGTGAAGTTCAGCAGGAG 233

Qy 337 k a l a g l u t h r a l a a s n l e u g i n a s p r o a s c i n l e u t y r a n g i u l e u a s n l e u g l 357
 |||::|
 Db 234 c c c a g a g c c c c c g c g t a c a g a g g g c c a g a a c a g c t t a t a a c g a g c t c a t c t a g 293

QY 357 YARGARGGLUGLUTyraspValleuGLuLysLARGALaARGaPProGLumetGLyGL 377
 DB 294 ACGAAGAGAGAGTACGATGTTTGGACAAGAGACGTGCCGGGACCCCTGAGATGGGGG 353

DQ 377 LYSGLNGINGARGARGASNPGRGNGIGLYValTYASNAlALEUGINLysAsPLY 397

DB 354 AAGCCCGAGAGNAGAACAACCTCAGGNAGCGCTGTCAATGAACTCGAGAAAGATTAA 413

QY 397 smetAAGIUAAtATyrserGIuIleGIYThrIrybGIyGIuArgaArgGIyLysGIyH 417
|||
Db 414 GATGGCGAGGCTCACTAGATTGGATTGAAGCGGAGCGCGGAGGGGCAAGGGGCA 473

QY 417 saspgyleurynglnglyseerTha.a.tnrlysasgThrlyrAsp.a.laleuhisme 437
|||
Db 474 CGATGCGCTTACAGGGTCTAGTAAGCCACCAAGAACCCTAAGCGGCCCTTCACAT 533

[illegible]

RESULT 13
LOCUS BM921412 1038 bp mRNA linear EST 12-MAR-2002
DEFINITION AGNCOURT_6626432 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5752844
5', mRNA sequence.
ACCESSION BM921412
VERSION BM921412.1 GI:19371791
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1038)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgsabbs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1AM12787 row: k column: 21
High quality sequence stop: 730.
Location/Qualifiers
1. 1038
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5752844"
/lab_host="DH10B"
/clone_id="NIH_MGC_115"
/note="Organ: pooled brain, lung, testis; Vector:
pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27, and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC library."

ORIGIN
Alignment Scores:
Pred. No.: 3,466-51 Length: 1038
Score: 620.00 Matches: 126
Percent Similarity: 79.64% Conservative: 7
Best Local Similarity: 75.45% Mismatches: 21
Query Match: 26.44% Indels: 13
DB: 5 Gaps: 2

US-09-596-774-7 (1-443) x BM921412 (1-1038)

QY 278 H1aBProThrglyThrSerGlnProGlnArgProGluAspCysArgProArgGlySerVal 297
Db 86 CATCTGAGGACAGCTTGGCCGATTACAGA----- 115
QY 298 LysGlyThrglyLeu-AspPheLeuGluAspProLysLeuGlyCysTyrLeuLeuAspGlyY11 317
Db 116 ---GGCACAAGAGCTTGGCTGCTG---GATCCCAAACTGCTACCTGCTGAGAGGAAT 169
QY 317 eleuPheIleTyrGlyValIleIleThrAlaLeuTyrLeuArgAlaLysPheSerArgSe 337
Db 170 CCTCTTCATCTATGTCATTCATTCACCTGCTTGTCTCTGAGAGGAAGTTCAGAGAG 229
QY 337 rAlaGluThrAlaAlaLeuLeuGlnAspProAsnGlnLeuTyrAsnGlnLeuAsnLeuG1 357
Db 230 CGCAGACGCCGCCGCTACAGAGGAGGACCAAGACCAAGCTATTAACGAGCTCAATCTAGG 289

QY 357 YArgArgGluGluTyrAspValLeuGluLysLysArgAlaArgAspProGluMetGlyG1 377
Db 290 ACGAAGAGAGAGTACGATGTTTGGACACAGACGTCGCCGGGACCTTGAGATGGGGG 349
QY 377 YLysGlnGlnArgArgArgAspProGlnGluGlyValTyrAsnAlaLeuGlnLysAspLy 397
Db 350 AAAGCCCGACAGAGGAAGAAACCTTCAGAGAGGCTGTACATATMACTGCAGAAAGATTA 409
QY 397 sMeAlaGluAlaTyrSerGluIleGlyThrLysGlyGluArgArgGlyLysGlyYH1 417
Db 410 GATGCGGAGGCTTACAGTGAATTTGGATGAAAGCGACGCGCGGAGGCGCAAGGCA 469
QY 417 sAspGlyLeuTyrGlnGlyLeuSerThrAlaThrLysAspThrTyrAsnAlaLeuH1sMe 437
Db 470 CGATGGCTTATACAGGGTCTCACTTACAGCCACAGAGACCTTACAGCCCTTACAT 529
QY 437 tGlnThrLeuAlaProArg 443
Db 530 GCAGCGCCCTGCGCCCTCGC 548

RESULT 14
LOCUS AL554350 1113 bp mRNA linear EST 30-MAR-2004
DEFINITION AL554350 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA
clone CSOD1082Y123 5-PRIME, mRNA sequence.
ACCESSION AL554350
VERSION AL554350.3 GI:45859108
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1113)
AUTHORS Li, W.B., Gruber, C., Jeese, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT On Feb 15, 2001 this sequence version replaced gi:31276162.
Genoscope - Centre National de Sequencage
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen. This sequence belongs to sequence cluster
484.x
For more information about this cluster, see
http://www.genoscope.cns.fr/cdnats=CSOD1082AE12QP1&c=484.r.

FEATURES
Source
Location/Qualifiers
1. 1113
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSOD1082Y123"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_id="Homo sapiens PLACENTA COT 25-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Alignment Scores:
Pred. No.: 3,82e-51 Length: 1113
Score: 620.00 Matches: 126
Percent Similarity: 79.64% Conservative: 7
Best Local Similarity: 75.45% Mismatches: 21
Query Match: 26.44% Indels: 13
DB: 1 Gaps: 2

US-09-596-774-7 (1-443) x AL554350 (1-1113)

```

Oy 278 HisProthrglyThrsSerGlnProGlnArgProGluAspCysArgProArgGlySerVal 297
Db CATCTGCGAGGACGATGTCGATTAACAG----- 116
Oy 298 LysGlyThrglyLeu-AspPheLeuGluAspProLysLeuCysTyrLeuLeuAspGlyI 317
Db 117 ---GGCAGACAGCTTGGCTGCTG---GATCCCAAACTGCTGCTGCTGCTGATGAT 170
Oy 317 eleupheiletyrGlyValIleIleThrsAlaLeuTyrLeuArgAlaLysPheSerArgSe 337
Db 171 CCTCTCATCTATGTCGATCTCTCATCTGCTGCTGCTGCTGAGAGGAACTTCAGAGAG 230
Oy 337 rAlaGluThrsAlaAlaAsnLeuGlnAspProAsnGlnLeuTyrAsnGlnLeuAsnLeuG 357
Db 221 CGCAACAGCCCCCGCTGACAGCAGAGGCGCAGAACAGCTCTATAACAGACTCAATCTAGG 290
Oy 357 YArgArgGluGluTyrAspValLeuGluLysLysArgAlaArgAspProGluMetGlyI 377
Db 291 ACGAAGAGAGAGTACGATGTTTGGACAAGAGAGCTGCGCGGACCTTGAGATGGGGG 350
Oy 377 YLysGlnGlnArgArgAsnProGlnGluGlyValTyrAsnAlaLeuGlnLysAspGly 397
Db 351 AAAGCCGAGAGAGAGAGAACCTTCAGAGAGGCTGTCATGAACTGCAGAAAGATGA 410
Oy 397 sMetAlaGluAlaTyrSerGluIleGlyThrsLysGlyLysArgArgGlyLysGlyI 417
Db 411 GATGCGGAGAGGCTACAGTGAATGGATGAAAGCGCAGCGGAGGCGCAAGGGGCA 470
Oy 417 sAspGlyLeuTyrGlnGlyLeuSerThrAlaThrLysAspThrTyrAspAlaLeuHis 437
Db 471 CGATGCGCTTTACAGGCTCTCAGTACAGCCACCAAGACCTTCAAGCGCTTCACAT 530
Oy 437 tGlnThrLeuAlaProArg 443
Db 531 GCAGGCGCTGCCCCCTGCG 549

RESULT 15
LOCUS CR625818 1546 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CS0DJ005YJ16 of T cells (Jurkat cell line)
ACCESSION CR625818
VERSION 1
KEYWORDS HTc; CNSLT_cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1546)
Full-length cDNA libraries and normalization
Unpublished
Contact: Feng Liang Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue
2 (bases 1 to 1546)
Genoscope.
Direct Submission
Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail: seqref@genoscope.cns.fr)
- Web : www.genoscope.cns.fr)
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
Location/Qualifiers
1 . 1546
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DJ005YJ16"
/tissue_type="T cells (Jurkat cell line) Cot

```

```

ORIGIN 10-normalized
/plasmid="pCMVSPORT_6"

Alignment Scores:
Pred. No.: 6.08e-51 Length: 1546
Score: 620.00 Matches: 126
Percent Similarity: 79.64% Conservative: 7
Best Local Similarity: 75.45% Mismatches: 21
Query Match: 26.44% Indels: 13
DB: 3 Gaps: 2

US-09-596-774-7 (1-443) x CR625818 (1-1546)
Oy 278 HisProthrglyThrsSerGlnProGlnArgProGluAspCysArgProArgGlySerVal 297
Db 69 CATCTGCGAGGACGATGTCGATTAACAG----- 98
Oy 298 LysGlyThrglyLeu-AspPheLeuGluAspProLysLeuCysTyrLeuLeuAspGlyI 317
Db 99 ---GGCAGACAGCTTGGCTGCTG---GATCCCAAACTGCTGCTGCTGATGAT 152
Oy 317 eleupheiletyrGlyValIleIleThrsAlaLeuTyrLeuArgAlaLysPheSerArgSe 337
Db 153 CCTCTCATCTATGTCGATCTCTCATCTGCTGCTGCTGCTGAGAGTGAAGTTCAGAGAG 212
Oy 337 rAlaGluThrsAlaAlaAsnLeuGlnAspProAsnGlnLeuTyrAsnGlnLeuAsnLeuG 357
Db 213 CGCAACAGCCCCCGCTGACAGCAGAGGCGCAGAACAGCTCTATAACAGACTCAATCTAGG 272
Oy 357 YArgArgGluGluTyrAspValLeuGluLysLysArgAlaArgAspProGluMetGlyI 377
Db 273 ACGAAGAGAGAGTACGATGTTTGGACAAGAGAGCTGCGCGGACCTTGAGATGGGGG 332
Oy 377 YLysGlnGlnArgArgAsnProGlnGluGlyValTyrAsnAlaLeuGlnLysAspGly 397
Db 333 AAAGCCGAGAGAGAGAGAACCTTCAGAGAGGCTGTCATGAACTGCAGAAAGATGA 392
Oy 397 sMetAlaGluAlaTyrSerGluIleGlyThrsLysGlyLysArgArgGlyLysGlyI 417
Db 393 GATGCGGAGAGGCTTACAGTGAATGGATGAAAGCGCAGCGGAGGCGCAAGGGGCA 452
Oy 417 sAspGlyLeuTyrGlnGlyLeuSerThrAlaThrLysAspThrTyrAspAlaLeuHis 437
Db 453 CGATGCGCTTTACAGGCTCTCAGTACAGCCACCAAGAGACACTTACAGCGCTTCACAT 512
Oy 437 tGlnThrLeuAlaProArg 443
Db 513 GCAGGCGCTGCCCCCTGCG 531

Search completed: April 20, 2005, 16:33:47
Job time : 4354.69 secs

```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2005, 06:50:16 ; Search time 63.2157 Seconds
(without alignments)
3588.524 Million cell updates/sec

Title: US-09-596-774-7
Perfect score: 2345
Sequence: 1 QIOLVSGPELKKRGGTETVKI.....LSTATKDTYDALHMQTLAPR 443

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt_03:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	904	38.6	241	2 Q921A6	Q921A6 mus musculu
2	832	35.5	244	2 Q65ZC8	Q65ZC8 homo sapien
3	778	33.2	240	2 Q65ZC9	Q65ZC9 homo sapien
4	776	33.1	255	2 Q6KB05	Q6KB05 mus musculu
5	767	32.7	243	2 Q7TQM2	Q7TQM2 mus musculu
6	746.5	31.8	218	2 Q925S1	Q925S1 mus musculu
7	744.5	31.7	298	2 Q9QYF0	Q9QYF0 synthetic c
8	733	31.3	487	2 Q65ZL2	Q65ZL2 mus sp. fv/
9	725.5	30.9	248	2 Q65Z07	Q65Z07 mus sp. b3/
10	713	30.4	164	1 CD3Z_MOUSE	P24161 mus musculu
11	618	26.4	164	1 CD3Z_HUMAN	P20963 homo sapien
12	605	25.8	206	1 CD3H_MOUSE	P23020 mus musculu
13	602.5	25.7	163	1 CD3Z_PIG	Q9X619 sus scrofa
14	589.5	25.1	165	1 CD3Z_RABIT	Q9UF68 oryctolagus
15	539	23.0	166	1 CD3Z_SHEEP	P23329 ovis aries
16	502.5	21.4	484	2 Q991A6	Q991A6 mus musculu
17	486	20.7	149	1 KV5A_MOUSE	P01633 mus musculu
18	454.5	19.4	116	2 Q683Y7	Q683Y7 mus musculu
19	446.5	19.0	140	2 Q65ZL3	Q65ZL3 mus sp. c91
20	429.5	18.3	116	2 Q683Y8	Q683Y8 mus musculu
21	419	17.9	102	2 Q91L79	Q91L79 mus musculu
22	417	17.8	108	2 Q8Y1U0	Q8Y1U0 mus musculu
23	413	17.6	99	2 Q9UL74	Q9UL74 mus musculu
24	409	17.4	170	2 Q925S2	Q925S2 mus musculu
25	398	17.0	125	2 Q6P1L0	Q6P1L0 homo sapien
26	394.5	16.8	518	2 Q6N030	Q6N030 homo sapien
27	393	16.8	497	2 Q8Y224	Q8Y224 homo sapien
28	391.5	16.7	236	2 Q6GMX8	Q6GMX8 homo sapien
29	390	16.6	136	1 KV5B_MOUSE	P01634 mus musculu
30	389	16.6	108	1 KV1Y_HUMAN	P80362 homo sapien
31	387.5	16.5	236	2 Q6P1H7	Q6P1H7 homo sapien

32	387.5	16.5	322	2 Q6KAV0	Q6KAV0 homo sapien
33	386.5	16.5	236	2 Q6GMW1	Q6GMW1 homo sapien
34	385.5	16.4	498	2 Q6N041	Q6N041 homo sapien
35	384	16.4	108	1 KV1Y_HUMAN	P04430 homo sapien
36	381.5	16.3	236	2 Q723YA	Q723YA homo sapien
37	380	16.2	480	2 Q6P089	Q6P089 homo sapien
38	379	16.2	108	1 KV1W_HUMAN	P01605 homo sapien
39	378	16.1	147	2 Q925S3	Q925S3 mus musculu
40	377.5	16.1	236	2 Q6GMX0	Q6GMX0 homo sapien
41	377	16.1	108	1 KV1B_HUMAN	P01594 homo sapien
42	377	16.1	108	1 KV1O_HUMAN	P01607 homo sapien
43	377	16.1	129	2 Q9UL77	Q9UL77 homo sapien
44	377	16.1	108	1 KV1W_HUMAN	P04431 homo sapien
45	375	16.0	108	2 Q9UL70	Q9UL70 homo sapien

ALIGNMENTS

RESULT 1

ID	Q921A6	PRELIMINARY:	PRT:	241 AA.
AC	Q921A6	01-DEC-2001 (TREMBLrel. 19, Created)		
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)			
DE	Anti-CEA 79 single chain Fv (Fragment).			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=98170165; PubMed=9509426;			
RA	Chung J.H., Choi S.J., Kim H.J., Kim I.J., Choi I.H., Lee S.D.,			
RA	Yi K.S., Suh P.G., Ryu S.H., Chung H.K.;			
RT	"Cloning and characterization of cDNAs encoding VH and VL of a			
RT	monoclonal anti-CEA antibody (CEA 79) cross-reactive with NCA-95 and			
RT	generation of a single-chain Fv molecule (scFv).";			
RL	Mol. Cells 7:816-819 (1997).			
DR	EMBL; U88067; AAB4804.1; -.			
DR	PIR; S19965; S19965.			
DR	PIR; S19967; S19967.			
DR	PIR; S19968; S19968.			
DR	PIR; S26325; S26325.			
DR	HSSP; P01607; 18MW.			
DR	SMART; SM00406; IGV; 2.			
DR	PROSITE; PS50835; IG_LIKE; 2.			
FT	NON_TER 1			
FT	NON_TER 241			
FT	NON_TER 241			
SQ	SEQUENCE 241 AA; 26086 MW; 0276887248B9C771 CRC64;			

Query Match 38.6%; Score 904; DB 2; Length 241;
Best Local Similarity 70.2%; Pred. No. 2.1e-58;
Matches 172; Conservative 29; Mismatches 34; Indels 10; Gaps 4;

QY	1 QIOLVSGPELKKRGGTETVKISCKASGYPTNNYNNWKAQAGGLKMMGWINTSGESTF 60		Q6KAV0	Q6KAV0
DB	1 QVRLQDSGPELKKRGGTETVKISCKASGYPTDYNNWKAQAGGLKMMGWINTTGEFTY 60		Q6GMW1	Q6GMW1
QY	61 ADDEKGRFDFSLSTSAATYALQINNLSKSDMATYFCARMEVYHGVYPWGCGTTVTVSSG 120		Q6N041	Q6N041
DB	61 ADDKGRFDFSLSTSAATYALQINNLSKSDMATYFCARKDLR-YFDVWGCGTTVTVSSG 119		KV1Y_HUMAN	KV1Y_HUMAN
QY	121 GGGSGGGSGGGSGSDIQLTQSHKFLSTSVGDRVSIITCKASQDVYNAVAWYQKKGSGPK- 179		Q723YA	Q723YA
DB	120 GGGSGGGSGGGSGSDIQLTQSHKFLSTSVGDRVSIITCKASQDVYNAVAWYQKKGSGPK- 179		Q6P089	Q6P089
QY	180 ---LLIYSASRYTVGVPSTFGSGSGPDPFTTSSVOAEDLAVYFCQGHPTPTFGSGT 236		Q925S3	Q925S3
DB	180 AHTLHYI---QGIPIRSRSGSGGRDYFSISNLEPEDLAIYCC-LHYDNLHTFGGCT 234		Q6GMX0	Q6GMX0
QY	237 KLEIK 241		KV1B_HUMAN	KV1B_HUMAN
DB			KV1O_HUMAN	KV1O_HUMAN
QY			Q9UL77	Q9UL77
DB			KV1W_HUMAN	KV1W_HUMAN
QY			Q9UL70	Q9UL70

Db 235 KLELK 239

RESULT 2

065ZC8 PRELIMINARY; PRT; 244 AA.

AC 065ZC8; 25-OCT-2004 (TREMBlrel. 28, Created)

DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)

DE Single-chain Fv (Fragment).

GN Name=scfv;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxId=9606;

RP [1]

RN SEQUENCE FROM N.A.

RX MEDLINE=97362799; PubMed=9219263;

RA Kontermann R.E., Wing M.G., Winter G.;

RT "Complement recruitment using bispecific diabodies."

RL Nat. Biotechnol. 15:629-631(1997).

DR EMBL; Y13057; CAA73500.1; -.

DR InterPro; IPR003599; Ig_.

DR InterPro; IPR007110; Ig_1like.

DR InterPro; IPR003596; Ig_v.

DR Pfam; PF00047; Ig; 2.

DR SMART; SM00409; IG; 2.

DR SMART; SM00406; IG; 2.

DR PROSITE; PSS0835; IG_LIKE; 2.

FT NON_TER 1

FT NON_TER 244

FT SEQUENCE 244 AA; 26127 MW; 4B1F17868338F2BF CRC64;

Query Match 35.5%; Score 832; DB 2; Length 244;

Best Local Similarity 61.9%; Pred. No. 4,1e-53;

Matches 151; Conservative 44; Mismatches 45; Indels 4; Gaps 2;

QY 1 QIOLVOSGPELKKRGETVVKISCKASGYPTNYGMNVWQAPGGGLKRMGMINTSGESTF 60

DB 1 QVOLVOSGAEVKKPKDSDVKVSCKASGYTFSDHYHVMVRQAPGGGLKRMGMINTSGESTF 60

QY 61 ADPRKGRFDSLETSANTAYLQINNLSKEDMATYFCAR--WEVYHGVVPYWGCTTVTS 117

DB 61 AQRFGRTVMTDRDISAAYMEVSRLESDDTAVVYCCAREGTSAYIG-MDVWGGTLVTV 119

QY 118 SSGGGSGGGSGGGSDIOLTQSHKFLSTSVGRVSIITCKASQDVYNAVWYQKPGSG 177

DB 120 SSGGGSGGGSGGGSDIOLTQSPSTLSASIGDRVITTCASGIGHMLAWYQKPGSKA 179

QY 178 PKLIYASRYTVGPSRFTSGSGSDPFTFTTSSVQADLAVYFCQGHFRPTFTFGSGTKL 237

DB 180 PKFLIYASRLASGAPSRFSSGSGTPTLTLSLQPDPAFTYCCQGYSNVPLTFGGGTR 239

QY 238 LEIK 241

DB 240 LEIK 243

RESULT 3

065ZC9 PRELIMINARY; PRT; 240 AA.

AC 065ZC9; 25-OCT-2004 (TREMBlrel. 28, Created)

DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)

DE Single-chain Fv (Fragment).

GN Name=scfv;

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxId=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C1G/77

RX MEDLINE=97362799; PubMed=9219263;

RA Kontermann R.E., Wing M.G., Winter G.;

RT "Complement recruitment using bispecific diabodies."

RL Nat. Biotechnol. 15:629-631(1997).

DR EMBL; Y13056; CAA73499.1; -.

DR InterPro; IPR003599; Ig_.

DR InterPro; IPR007110; Ig_1like.

DR InterPro; IPR003596; Ig_v.

DR Pfam; PF00047; Ig; 2.

DR SMART; SM00409; IG; 2.

DR SMART; SM00406; IG; 2.

DR PROSITE; PSS0835; IG_LIKE; 2.

FT NON_TER 1

FT NON_TER 240

FT SEQUENCE 240 AA; 25569 MW; FDCFD345F64B373 CRC64;

Query Match 33.2%; Score 778; DB 2; Length 240;

Best Local Similarity 60.3%; Pred. No. 3.7e-49;

Matches 146; Conservative 39; Mismatches 53; Indels 4; Gaps 2;

QY 1 QIOLVOSGPELKKRGETVVKISCKASGYPTNYGMNVWQAPGGGLKRMGMINTSGESTF 60

DB 1 QVOLVOSGGGLVPGGSLRSCASGFTSSYGMHVMVRQAPGGGLKRMGMINTSGESTF 60

QY 61 ADPRKGRFDSLETSANTAYLQINNLSKEDMATYFCAR--WEVYHGVVPYWGCTTVTS 119

DB 61 ADSVKGRTFISRNDSKNTLYLQNNLSRAEDTAVVYCCARDW--GSDLPDWGKCTLVTS 117

QY 120 GGGSGGGSGGGSGSDIOLTQSHKFLSTSVGRVSIITCKASQDVYNAVWYQKPGSPK 179

DB 118 GGGSGGGSGGGSGSDIOLTQSPSTLSASIGDRVITTCASGIGHMLAWYQKPGSKA 177

QY 180 LLIYASRYTVGPSRFTSGSGSDPFTFTTSSVQADLAVYFCQGHFRPTFTFGSGTKLE 239

DB 178 LLIYKASRLASRAPSRFSSGSGTPTLTLSLQPDPAFTYCCQGYSNVPLTFGGGTR 237

QY 240 IK 241

DB 238 IK 239

RESULT 4

06KB05 PRELIMINARY; PRT; 255 AA.

AC 06KB05; 05-JUL-2004 (TREMBlrel. 27, Created)

DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)

DE scfv B8E5 protein (Fragment).

GN Name=scfv B8E5.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxId=10090;

RP [1]

RN SEQUENCE FROM N.A.

RA STRAIN=Bald/c;

RC Peter U.C., Wallukat G., Tugler J., Maurice D., Roegel J.C.,

RA Briand J.P., Hoebeke J.;

RL Submitted (JUN-2004) to the EMBL/Genbank/DBJ databases.

DR EMBL; AJ746180; CAG34081.1; -.

DR HSSP; P01837; 1KCR.

DR InterPro; IPR003599; Ig_.

DR InterPro; IPR007110; Ig_1like.

DR InterPro; IPR003596; Ig_v.

DR SMART; SM00409; IG; 2.

DR SMART; SM00406; IG; 2.

DR PROSITE; PSS0835; IG_LIKE; 2.

FT NON_TER 1

FT NON_TER 255

FT SEQUENCE 255 AA; 27445 MW; B68BD38395DF713B CRC64;

Query Match 33.1%; Score 776; DB 2; Length 255;
Best Local Similarity 59.8%; Pred. No. 5.7e-49;
Matches 149; Conservative 36; Mismatches 56; Indels 8; Gaps 2;

QY 1 QIOLVSGPELKKPEETVYKISCKASGYPTNYGMNWKQAPEGQKMMGMINSTSGESTF 60
DB 1 QVQLQSGSGDIYVKPGSGLKVKSCAASGFTFSSYGMWVRQTPDKRLAEVATITSGGSYTY 60
QY 61 ADDEKGRFDSLETSANTAYLQINNLKSEDMATYFCARMEVYHGVPYWGQGITVTVSSG 118
DB 61 PDSVAGRTTIRSDNKNLTLYQMSLSKSEDTAMVYCAHINRYDGAFDYWGQGITVTVSS 120
QY 119 SGGSGSGSGSGSGSDIQLTQSHKFLSTVSDRVYSITCKASQDVYNA-----VAATYQQ 172
DB 121 SGGSGSGSGSGSGSDIYVMAQPSLSVSAAGEKYMCKSSQSLNLRNKNYLAATYQQ 180
QY 173 KPGQSPKLLIYASASRYTGVPSPFTGSGSDPFTFTISSVQAEDLAVYPCQHPFTPTPT 232
DB 181 KPGQSPKLLIYASASRYTGVPSPFTGSGSDPFTFTISSVQAEDLAVYPCQHPFTPTPT 240
QY 233 GSGTKLEIK 241
DB 241 GAGTKLEIK 249

RESULT 5

Q7TQM2 PRELIMINARY; PRT; 243 AA.
AC Q7TQM2;
DT 01-OCT-2003 (TREMBlrel. 25, Created)
DT 01-OCT-2003 (TREMBlrel. 25, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Scfv 6H8 protein (Fragment).
GN Name=Scfv 6H8;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RX MEDLINE=22853226; PubMed=12860977; DOI=10.1074/jbc.M306877200;
RA Peter J.C., Eftekhari P., Billiard P., Wallukke G., Hoebeke J.;
RT "scfv single chain antibody variable fragment as inverse agonist for
the beta-2 adrenergic receptor."
RL J. Biol. Chem. 278:36740-36747(2003).
DR EMBL; AJ574851; CAE00495.1; -.
DR HSBP; P01751; IAGW.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003596; IG_v.
DR SMART; SM00406; IGV; 2.
DR PROSITE; PS50835; IG-LIKE; 2.
FT NON TER 1
FT PROSITE 1
SQ SEQUENCE 243 AA; 25976 MW; BEFF642DCFAF76 CRC64;

Query Match 32.7%; Score 767; DB 2; Length 243;
Best Local Similarity 60.6%; Pred. No. 2.4e-48;
Matches 146; Conservative 42; Mismatches 49; Indels 4; Gaps 3;

QY 1 QIOLVSGPELKKPEETVYKISCKASGYPTNYGMNWKQAPEGQKMMGMINSTSGESTF 60
DB 1 QVQLQSGSELVPRGASVKLSCKASGYFTTYMMWVAQRHGQGLEMINIYPPSGITNY 60
QY 61 ADDEKGRFDSLETSANTAYLQINNLKSEDMATYFCARMEVYHGVPYWGQGITVTVSSG 120
DB 61 DEKRNKILTVDTSSSTAYVHLSLASEDSAVVYCARGG--RG-LDVGAGTTLTVSSG 117
QY 121 GGGSGSGSGSGSDIQLTQSHKFLSTVSDRVYSITCKASQDVYNAVAATYQQKQSPKL 180
DB 118 GGGSGSGSGSGSDIQLTQSHKFLSTVSDRVYSITCKASQDVYNAVAATYQQKQSPKL 177
QY 181 LIYASASRYTGVPSPFTGSGSDPFTFTISSVQAEDLAVYPCQHPFTPTPTFGSGTKLEI 240

DB 178 LISGATSIETGVSPRFGSGSGKDYTLSTLQTEDVATYVCOQWSTR-TFGGSGTKLEI 236
QY 241 K 241
DB 237 K 237

RESULT 6

Q925S1 PRELIMINARY; PRT; 218 AA.
AC Q925S1;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE MRES (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RX PubMed=11819679;
RA Cui D., Zeng G., Yan X., Wang F., Tian F., Ren D., Zhao T., Li X.,
RA Su C.;
RT "Mechanism of exogenous nucleic acids and their precursors improving
the repair of intestinal epithelium after irradiation in mice."
RL World J. Gastroenterol. 6:709-717(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RA Cui D., Zeng G., Yan X., Li X., Su C.;
RT "Cloning of mouse genes related to repairing of intestinal epithelium
of the irradiated mice by treatment with the intestinal RNA of mice of
the same strain."
RL Int. J. Radiat. Biol. Relat. Stud. Phys. Chem. Med. 19:71-80(2001).
DR EMBL; AF240168; AAK43733.1; -.
DR HSBP; P01665; IONZ.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG-LIKE; 1.
FT NON TER 218
SQ SEQUENCE 218 AA; 23013 MW; 527E4FA8F7982817 CRC64;

Query Match 31.8%; Score 746.5; DB 2; Length 218;
Best Local Similarity 66.2%; Pred. No. 6.8e-47;
Matches 141; Conservative 28; Mismatches 39; Indels 5; Gaps 2;

QY 1 QIOLVSGPELKKPEETVYKISCKASGYPTNYGMNWKQAPEGQKMMGMINSTSGESTF 60
DB 3 QVQLQSGPELKKPEETVYKISCKASGYFTTAKQWVAQKPGKLMKMGMINSTSGVPKY 62
QY 61 ADDEKGRFDSLETSANTAYLQINNLKSEDMATYFCARMEVYHGVPYWGQGITVTVSSG 120
DB 63 ABEFKGRFARSLTSASTAYLQISNLKNEDTATFCRWD-YDGGFAYWGQGITVTVSSG 121
QY 121 GGGSGSGSGSGSDIQLTQSHKFLSTVSDRVYSITCKASQDVYNA-----VAATYQQKQ 176
DB 122 GGGSGSGSGSGSDIQLTQSHKFLSTVSDRVYSITCKASQDVYNA-----VAATYQQKQ 181
QY 177 SPKLLIYASASRYTGVPSPFTGSGSDPFTFTISSVQAEDLAVYPCQHPFTPTPTFGSGTKLEI 240
DB 182 PKLLIYASASRYTGVPSPFTGSGSDPFTFTISSVQAEDLAVYPCQHPFTPTPTFGSGTKLEI 240

RESULT 7

Q9QYF0 PRELIMINARY; PRT; 298 AA.
AC Q9QYF0;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE CN 8 single chain antibody.

DB 3 VKLVESGGGLVPGGSLKSLSCATSGFTSDYVYMWVROTPEKRLMEVAVYISNDSSAAYS 62
 QY 62 DDGGRPFSLSTANTAYLQINNKSRDMATPCARMEVHYGVYWGQGTTLTVSSGG 121
 DB 63 DTVGKRFITSDNANNTLYLQWRSRKSEDTALYSARGILAMGAWAYWGQGLTVVSSGG 122
 QY 122 GSGGGGGGGGGGDIQTLQSHKFLSTSVQDRVSTTCASODPVNA-----VAMVQOQKPGQ 176
 DB 123 GSGGGGGGGGGGSDVLTMTQSPSLFVSLGDDASISCRSQIIIVHSNGMTYLEWYLQKPGQ 182
 QY 177 SPKLLIYSASSRYGVPSRFRTGSGGPPFTFTTSSVQADLAVYFCQGHFTPTFGSGT 236
 DB 183 SPKLLIYKVSNNRFGVDPDRFSGSGGTFTLKISRVEAEDLGVVYCFQGSHPFTFGSGT 242
 QY 237 KLEIKR 242
 DB 243 KLEIKR 248
 RESULT 10
 CD3Z_MOUSE STANDARD: PRT: 164 AA.
 ID CD3Z_MOUSE
 AC P24161; Q9D3G3;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE T-cell surface glycoprotein CD3 zeta chain precursor (T-cell receptor
 T3 zeta chain).
 GN Name=CD3z; Synonyms=Tcrz;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE=86145643; PubMed=3278377;
 RA Weissman A.M., Banayash M., Hou D., Samelson L.E., Burgess W.H.,
 RA Klausner R.D.;
 RT "Molecular cloning of the zeta chain of the T cell antigen receptor.";
 RL Science 239:1018-1021 (1988).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Liver;
 RC MEDLINE=89327299; PubMed=2787796;
 RA Banayash M., Hsu V.W., Seldin M.F., Klausner R.D.;
 RT "The isolation and characterization of the murine T cell antigen
 receptor zeta chain gene.";
 RL J. Biol. Chem. 264:13252-13257 (1989).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Thymus;
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
 RA Okazaki Y., Futuro M., Kasukawa T., Adachi J., Bono H., Kondo S.,
 RA Nakajima I., Osato N., Saito R., Suzuki H., Yamanaoka I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gotohori T.,
 RA Beldick R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schirral L.M., Kanapin A., Matwuda H., Batalov S., Beisel K.W.,
 RA Blake J.A., Brad D., Brusic V., Chochia C., Corbani L.E., Cousins S.,
 RA Dalla E., Drisgani T.A., Fletcher C.F., Forrest A., Gough J.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimonard S., Guslich S., Hirokawa N., Jackson I.J., Jarvis E.D.,
 RA Kanai A., Kawanji H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Petrea G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramchandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Sempke C.A., Setou M., Shimada K.,
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Vitarito R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wyshew-Borls A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carlini P., Hayatsu N.,
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,

RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Imocani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,
 RA Yaenunthi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573 (2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6; TISSUE=Hematopoietic;
 RX MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ueclin T.B., Toshlyuk S., Carlini P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,
 RA Villalón D.K., Wuzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Maiza M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [5]
 RP PARTIAL SEQUENCE.
 RX MEDLINE=90239005; PubMed=2139725;
 RA Jin Y.J., Clayton L.K., Howard F.D., Koyasu S., Stieh M.,
 RA Steinbrich R., Tarr G.B., Reinherz E.L.;
 RT "Molecular cloning of the CD3 zeta subunit identifies a CD3 zeta-
 RT related product in thymus-derived cells.";
 RL Proc. Natl. Acad. Sci. U.S.A. 87:3319-3323 (1990).
 RN [6]
 RP ALTERNATIVE SPLICING.
 RX PubMed=2150596;
 RA Ohno H., Saito T.;
 RT "CD3 zeta and eta chains are produced by alternative splicing from a
 RT common gene.";
 RL Int. Immunol. 2:1117-1119 (1990).
 RN [7]
 RP ERRATUM.
 RA Ohno H., Saito T.;
 RL Int. Immunol. 4:1339-1339 (1992).
 RN [8]
 RP INTERACTION WITH SLA.
 RX MEDLINE=20130290; PubMed=10662792;
 RA Sosinowski T., Pandey A., Dixit V.M., Weiss A.;
 RT "Src-like adaptor protein (SLAP) is a negative regulator of T cell
 RT receptor signaling.";
 RL J. Exp. Med. 191:463-474 (2000).
 RN [9]
 RP INTERACTION WITH SLA2.
 RX MEDLINE=22011997; PubMed=11891219; DOI=10.1074/jbc.M110318200;
 RA Pandey A., Ibarola N., Kratchmarova I., Fernandez M.M.,
 RA Constantinescu S.N., Chao O., Sawadikoso S., Lodish H.F., Mann M.,
 RA "A novel Src homology 2 domain-containing molecule, Src-like adapter
 RT protein-2 (SLAP-2), which negatively regulates T cell receptor
 RT signaling.";
 RL J. Biol. Chem. 277:19131-19138 (2002).
 CC -I- FUNCTION: Probable role in assembly and expression of the TCR
 CC complex as well as signal transduction upon antigen triggering.
 CC -I- SUBUNIT: Interacts with DOK2 (by similarity). The TCR/CD3 complex
 CC of T lymphocytes consists of either a TCR alpha/beta or TCR
 CC gamma/delta heterodimer coexpressed at the cell surface with the
 CC invariant subunits of CD3 labeled gamma, delta, epsilon, zeta, and

CC eta. CD3-zeta forms either homodimers or heterodimers with CD3-
 CC eta. Interacts with SLA and SLA2.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=CD-3-zeta;
 CC IsoId=P24161-1; Sequence=Displayed;
 CC Name=CD-3-eta;
 CC IsoId=P29020-1; Sequence=External;
 CC -1- PTM: Phosphorylated on Tyr residues after T-cell receptor
 CC triggering.
 CC -1- SIMILARITY: Belongs to the CD3Z/FCER1G family.
 CC -1- SIMILARITY: Contains 3 ITAM domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M19729; AAA40171.1; -;
 CC EMBL: J04967; AAA50301.1; -;
 CC EMBL: AK017904; BAB30997.1; -;
 CC EMBL: BC052824; AAH52824.1; -;
 CC PIR: A40104; A40104.
 CC MGD: MGI:88334; Cd3z.
 CC InterPro: IPR003110; ITAM.
 CC Pfam: PF02189; ITAM; 3.
 CC Alternative splicing; Direct protein sequencing; Phosphorylation;
 CC Receptor; Repeat; Signal; T-cell; Transmembrane.
 CC KW
 CC FT
 CC FT CHAIN 1 21
 CC FT SIGNAL 1 21
 CC FT CHAIN 22 164
 CC FT T-cell surface glycoprotein CD3 zeta
 CC FT chain.
 CC FT DOMAIN 22 30 Extracellular (Potential).
 CC FT TRANSMEM 31 51 Potential.
 CC FT DOMAIN 52 164 Cytoplasmic (Potential).
 CC FT DOMAIN 69 89 ITAM 1.
 CC FT DOMAIN 108 129 ITAM 2.
 CC FT DOMAIN 139 159 ITAM 3.
 CC FT DISULFID 32 32 Interchain (Potential).
 CC FT MOD RSS 153 153 Phosphotyrosine.
 CC FT CONFLICT 153 153 Y->C (in Ref. 3).
 CC FT SEQUENCE 164 AA; 18637 MW; 1B8022035A312831 CRC64;
 CC SQ
 CC Query Match 30.4%; Score 713; DB 1; Length 164;
 CC Best Local Similarity 99.3%; Pred. No. 1.4e-44;
 CC Matches 138; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 CC
 CC QY 305 LEDPKLCYLDGILFYGVITITALLYLPAKFSRAETANLQDPQVLYNEINLGRREBYDV 364
 CC DB 26 LLDPKLCYLDGILFYGVITITALLYLPAKFSRAETANLQDPQVLYNEINLGRREBYDV 85
 CC QY 365 LEKRRADPEMGKQORRRNPQEGVNALQDKVAEYSEITGGERRRKGHDGLYQGL 424
 CC DB 86 LEKRRADPEMGKQORRRNPQEGVNALQDKVAEYSEITGGERRRKGHDGLYQGL 145
 CC QY 425 STATKQTYDALHMQTLAPR 443
 CC DB 146 STATKQTYDALHMQTLAPR 164
 CC
 CC RESULT 11
 CC ID CD3Z HUMAN STANDARD; PRT; 164 AA.
 CC AC P20963; OBTAX4;
 CC DT 01-FEB-1991 (Rel. 17, Created)
 CC DT 28-FEB-2003 (Rel. 41, Last sequence update)
 CC DT 25-OCT-2004 (Rel. 45, Last annotation update)
 CC DE T-cell surface glycoprotein CD3 zeta chain precursor (T-cell receptor
 CC T3 zeta chain).
 CC GN Name=CD3Z; Synonyms=T3Z, TCRZ;

OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 CC NCBI_TaxId=9606;
 CC
 CC RN
 CC RP SEQUENCE FROM N.A.
 CC RX MEDLINE=89071765; PubMed=2974162;
 CC RA Weissman A.M., Hou D., Orloff D.G., Modi W.S., Seunanez H.,
 CC O'Brien S.J., Klausner R.D.;
 CC RT "Molecular cloning and chromosomal localization of the human T-cell
 CC receptor zeta chain: distinction from the molecular CD3 complex.";
 CC Proc. Natl. Acad. Sci. U.S.A. 85:9709-9713(1988).
 CC [2]
 CC RN
 CC RP SEQUENCE FROM N.A.
 CC RX TISSUE=Pancreas;
 CC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 CC RA Strausberg R.L., Feingold E.A., Grouse L.H., Dege J.G.,
 CC Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 CC Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
 CC Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heist F.,
 CC Diatchenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,
 CC Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
 CC Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
 CC Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 CC Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 CC Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 CC Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 CC Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 CC Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 CC Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 CC Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 CC Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 CC Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 CC "Generation and initial analysis of more than 15,000 full-length human
 CC and mouse cDNA sequences.";
 CC Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC [3]
 CC RN
 CC RP INTERACTION WITH HIV-1 NEF.
 CC RX PubMed=10224289;
 CC RA Xu X.-N., Laferriere B., Srean G.R., Kraft M., Wolf D., Kolanus W.,
 CC Mongkolsapay J., McMichael A.J., Baur A.S.;
 CC RT "Induction of Fas ligand expression by HIV involves the interaction of
 CC Nef with the T cell receptor zeta chain.";
 CC J. Exp. Med. 189:1489-1496(1999).
 CC [4]
 CC RN
 CC RP INTERACTION WITH SLA.
 CC RX MEDLINE=99380595; PubMed=10449770; DOI=10.1073/pnas.96.17.9775;
 CC RA Tang J., Sawaadikogol S., Chang J.-H., Burakoff S.J.;
 CC RT "SLAP, a dimeric adapter protein, plays a functional role in T cell
 CC receptor signaling.";
 CC Proc. Natl. Acad. Sci. U.S.A. 96:9775-9780(1999).
 CC [5]
 CC RN
 CC RP INTERACTION WITH DOCK2.
 CC RX MEDLINE=22165501; PubMed=12176041; DOI=10.1016/S0006-291X(02)00931-2;
 CC RA Nishihara H., Maeda M., Tsuda M., Makino Y., Sawa H., Nagashima K.,
 CC Tanaka S.;
 CC RT "DOCK2 mediates T cell receptor-induced activation of Rac2 and IL-2
 CC transcription.";
 CC Biochem. Biophys. Res. Commun. 296:716-720(2002).
 CC [6]
 CC RN
 CC RP PHOSPHORYLATION SITES TYR-83; TYR-111; TYR-123 AND TYR-142.
 CC RX PubMed=12522270; DOI=10.1073/pnas.2436191100;
 CC RA Salomon A.R., Picardo S.B., Brill L.M., Brinker A., Phung Q.T.,
 CC Ericson C., Sauer K., Brock A., Horn D.M., Schultz P.G., Peters E.C.;
 CC RT "Profiling of tyrosine phosphorylation pathways in human cells using
 CC mass spectrometry.";
 CC Proc. Natl. Acad. Sci. U.S.A. 100:443-448(2003).
 CC [7]
 CC RN
 CC RP STRUCTURE BY NMR OF 136-149.
 CC RX MEDLINE=93201600; PubMed=7680960; DOI=10.1016/0092-8674(93)90405-F;
 CC RA Wakeman G., Shoelson S.E., Pant N., Cowburn D., Kurtyan J.;
 CC RT "Binding of a high affinity phosphotyrosyl peptide to the Src SH2
 CC domain: crystal structures of the complexed and peptide-free forms.";

```

RL Cell 72:779-790(1993).
CC -1- FUNCTION: Probable role in assembly and expression of the TCR
CC complex as well as signal transduction upon antigen triggering.
CC -1- SUBUNIT: The TCR/CD3 complex of T lymphocytes consists of either a
CC TCR alpha/beta or TCR gamma/delta heterodimer coexpressed at the
CC cell surface with the invariant subunits of CD3 labeled gamma,
CC delta, epsilon, zeta, and eta. CD3-zeta forms either homodimers or
CC heterodimers with CD3-eta. Interacts with SIA and SIA2. Interacts
CC with DOCK2. Interacts with HIV-1 Nef protein.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=CD-3-zeta;
CC IsoId=P20963-1; Sequence=Displayed;
CC Name=CD-3-eta;
CC IsoId=P20963-2; Sequence=Not described;
CC -1- PTM: Phosphorylated on Tyr residues after T-cell receptor
CC triggering (By similarity).
CC -1- SIMILARITY: Belongs to the CD32/FCER1G family.
CC -1- SIMILARITY: Contains 3 ITAM domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J04132; AAA60394.1; -.
DR EMBL: BC025703; AAH25703.1; -.
DR PIR: A31768; A31768.
DR PDB: 1TCE; NMR: B-136-149.
DR GeneW: HGNC:1677; CD3Z.
DR H-InvDB: HIX0001296; -.
DR MIM: 186780; -.
DR GO: GO:0005886; C:Plasma membrane; TAS.
DR GO: GO:0042101; C:T-cell receptor complex; TAS.
DR GO: GO:0005515; F:protein binding; IPI.
DR GO: GO:0042803; F:protein homodimerization activity; NAS.
DR InterPro: IPR003110; ITAM.
DR Pfam: PF02189; ITAM; 3.
DR SMART: SM00077; ITAM; 3.
DR 3D-structure: Alternative splicing; Phosphorylation; Receptor; Repeat;
KM Signal; T-cell; Transmembrane.
FT SIGNAL 1 21
FT CHAIN 22 164
FT DOMAIN 22 30 T-cell surface glycoprotein CD3 zeta
FT TRANSSEM 31 51 chain.
FT DOMAIN 52 164 Extracellular (Potential).
FT DOMAIN 69 89 Potential.
FT DOMAIN 108 129 Cytoplasmic (Potential).
FT DOMAIN 139 159 ITAM 1.
FT DISUPID 32 32 ITAM 2.
FT MOD_RES 83 83 ITAM 3.
FT MOD_RES 83 83 Interchain (Potential).
FT MOD_RES 111 111 Phosphotyrosine.
FT MOD_RES 123 123 Phosphotyrosine.
FT MOD_RES 142 142 Phosphotyrosine.
FT MOD_RES 153 153 Phosphotyrosine (By similarity).
FT CONFLICT 60 61 DA -> EP (in Ref. 1).
FT CONFLICT 101 101 Missing (in Ref. 1).
FT SEQUENCE 164 AA; 18696 MW; 940826034856EE9 CMC64;

Query Match 26.4%; Score 618; DB 1; Length 164;
Best Local Similarity 84.9%; Pred. No. 1.3e-37;
Matches 118; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

```

```

DB 86 LDKRRGDPENWGKQRRRNPQEGVNALQDKMAEAYSEIGTGERRRGKHDGLYQGL 145
QY 425 STATKQTYDALHNOTLAPR 443
DB 146 STATKQTYDALHNOTLAPR 164

RESULT 12
CD3H_MOUSE STANDARD; PRT; 206 AA.
AC P29020;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE T-cell surface glycoprotein CD3 eta chain precursor (T-cell receptor
DE T3 eta chain).
GN Name=Cd3z; Synonyms=Cd3h;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=90239005; PubMed=2139725;
RA Jin Y.J., Clayton L.K., Howard F.D., Koyasu S., Sieh M.,
RA Reinherz E.L., Tarr G.E., Reinherz E.L.;
RT "Molecular cloning of the CD3 eta subunit identifies a CD3 zeta-
RT related product in thymus-derived cells."
RT Proc. Natl. Acad. Sci. U.S.A. 87:3319-3323(1990).
RN [2]
RP SEQUENCE OF 144-206 FROM N.A.
RX MEDLINE=91271358; PubMed=1828894;
RA Clayton L.K., D'Adamo L., Sieh M., Hussey R.E., Koyasu S.,
RA Reinherz E.L., Howard F.B.;
RT "CD3 eta and CD3 zeta are alternatively spliced products of a common
RT genetic locus and are transcriptionally and/or post-transcriptionally
RT regulated during T-cell development."
RT Proc. Natl. Acad. Sci. U.S.A. 88:5202-5206(1991).
RN [3]
RP SEQUENCE OF 144-206 FROM N.A.
RX PubMed=2150596;
RA Ohno H., Saito T.;
RT "CD3 zeta and eta chains are produced by alternative splicing from a
RT common gene."
RT Int. Immunol. 2:1117-1119(1990).
RN [4]
RP ERRATUM.
RA Ohno H., Saito T.;
RL Int. Immunol. 4:1339-1339(1992).
CC -1- FUNCTION: Probable role in assembly and expression of the TCR
CC complex as well as signal transduction upon antigen triggering.
CC -1- SUBUNIT: The TCR/CD3 complex of T lymphocytes consists of either a
CC TCR alpha/beta or TCR gamma/delta heterodimer coexpressed at the
CC cell surface with the invariant subunits of CD3 labeled gamma,
CC delta, epsilon, zeta, and eta. CD3-eta can be complexed in a
CC heterodimeric form with CD3-zeta subunit. CD3-eta homodimer has
CC not been observed.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=CD-3-eta;
CC IsoId=P29020-1; Sequence=Displayed;
CC Name=CD-3-zeta;
CC IsoId=P29020-1; Sequence=External;
CC -1- SIMILARITY: Belongs to the CD32/FCER1G family.
CC -1- SIMILARITY: Contains 3 ITAM domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

entitles requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; M33158; AAA37398.1; -
 DR EMBL; M76711; AAA40403.1; -
 DR PIR; A35900; A35900.
 DR MGI; MGI:88334; Cd3z.
 DR InterPro; IPR003110; ITAM.
 DR Pfam; PF02189; ITAM; 2.
 DR SMART; SM00077; ITAM; 2.
 DR SMART; SM00077; ITAM; 2.
 KM Alternative splicing; Direct protein sequencing; Receptor; Repeat;
 FT Signal; T-cell; Transmembrane.
 FT CHAIN 1 21
 FT SIGNAL 1 21
 FT CHAIN 22 206
 FT TRANSMEM 22 30
 FT DOMAIN 31 51
 FT DOMAIN 52 206
 FT DOMAIN 69 89
 FT DOMAIN 108 129
 FT DOMAIN 139 159
 FT DISULFID 32 32
 FT DISULFID 32 32
 SO SEQUENCE 206 AA; 23339 MW; 829256A2CF444444 CRC64;

Query Match 25.8%; Score 605; DB 1; Length 206;
 Best Local Similarity 99.2%; Pred. No. 1.6e-36;
 Matches 117; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 305 LEDPKCYLDDGLFTYGVITITLYLRKFSRAETANLQDPNQLYNELNGRREYDV 364
 DB 26 LDDPKCYLDDGLFTYGVITITLYLRKFSRAETANLQDPNQLYNELNGRREYDV 85

QY 365 LEKRRARDPEMGKQOORRRNPQEGVYNALQKDKMAEAYSEIGTGERRRGKHGDLGYQ 422
 DB 86 LEKRRARDPEMGKQOORRRNPQEGVYNALQKDKMAEAYSEIGTGERRRGKHGDLGYQ 143

RESULT 13
 CD32_PIG STANDARD; PRT; 163 AA.
 AC O9XST9;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE T-cell surface glycoprotein CD3 zeta chain precursor (T-cell receptor T3 zeta chain).
 GN Name=CD32;
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
 OC NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Minnesota miniature swine;
 RA Jie H.-B., Yim D., Kim Y.B.;
 RT "The molecular cloning of porcine CD3 zeta.";
 RL Submitted (MAY-1999) to the EMBL/Genbank/DBJ databases.
 CC -I- FUNCTION: Probable role in assembly and expression of the TCR complex as well as signal transduction upon antigen triggering.
 CC -I- SUBUNIT: The TCR/CD3 complex of T lymphocytes consists of either a TCR alpha/beta or TCR gamma/delta heterodimer coexpressed at the cell surface with the invariant subunits of CD3 labeled gamma, delta, epsilon, zeta, and eta (By similarity).
 CC -I- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -I- PTM: Phosphorylated on Tyr residues after T-cell receptor triggering (By similarity).
 CC -I- SIMILARITY: Belongs to the CD32/FCER1G family.
 CC -I- SIMILARITY: Contains 3 ITAM domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; AF153830; AAD34640.1; -
 DR InterPro; IPR003110; ITAM.
 DR Pfam; PF02189; ITAM; 3.
 DR SMART; SM00077; ITAM; 3.
 KM Phosphorylation; Receptor; Repeat; Signal; T-cell; Transmembrane.
 FT Signal 1 21
 FT CHAIN 22 163
 FT SIGNAL 22 163
 FT CHAIN 22 163
 FT TRANSMEM 31 51
 FT DOMAIN 52 163
 FT DOMAIN 69 89
 FT DOMAIN 107 128
 FT DOMAIN 138 158
 FT DISULFID 32 32
 FT MOD RES 152 152
 FT MOD RES 152 152
 SO SEQUENCE 163 AA; 18566 MW; 34898620B67167C7 CRC64;

Query Match 25.7%; Score 602.5; DB 1; Length 163;
 Best Local Similarity 84.2%; Pred. No. 1.8e-36;
 Matches 117; Conservative 7; Mismatches 14; Indels 1; Gaps 1;

QY 305 LEDPKCYLDDGLFTYGVITITLYLRKFSRAETANLQDPNQLYNELNGRREYDV 364
 DB 26 LDDPKCYLDDGLFTYGVITITLYLRKFSRAETANLQDPNQLYNELNGRREYDV 85

QY 365 LEKRRARDPEMGKQOORRRNPQEGVYNALQKDKMAEAYSEIGTGERRRGKHGDLGYQ 424
 DB 86 LDRRRGSDPEMGK -PRRNPQEGVYNALQKDKMAEAYSEIGTGERRRGKHGDLGYQ 144

QY 425 STATKDTYDALHMQTLAPR 443
 DB 145 STATKDTYDALHMQTLAPR 163

RESULT 14
 CD32_RABIT STANDARD; PRT; 165 AA.
 AC Q9TUF8;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE T-cell surface glycoprotein CD3 zeta chain precursor (T-cell receptor T3 zeta chain).
 GN Name=CD32;
 OS Oryctolagus cuniculus (Rabbit).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
 OC NCBI_TaxID=9986;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=B/J X Chbb.HM;
 RA Isono T., Nishimura M.;
 RT "Rabbit CD3 zeta.";
 RL Submitted (NOV-1999) to the EMBL/Genbank/DBJ databases.
 CC -I- FUNCTION: Probable role in assembly and expression of the TCR complex as well as signal transduction upon antigen triggering.
 CC -I- SUBUNIT: The TCR/CD3 complex of T lymphocytes consists of either a TCR alpha/beta or TCR gamma/delta heterodimer coexpressed at the cell surface with the invariant subunits of CD3 labeled gamma, delta, epsilon, zeta, and eta (By similarity).
 CC -I- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -I- PTM: Phosphorylated on Tyr residues after T-cell receptor triggering (By similarity).
 CC -I- SIMILARITY: Belongs to the CD32/FCER1G family.
 CC -I- SIMILARITY: Contains 3 ITAM domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; ABO35152; BAA6994.1; -
 DR InterPro; IPR003110; ITAM.
 DR Pfam; PF02189; ITAM; 3.
 DR SMART; SM00077; ITAM; 3.
 KW Phosphorylation; Receptor; Repeat; Signal; T-cell; Transmembrane.
 FT SIGNAL 1 21 By similarity.
 FT CHAIN 22 165 T-cell surface glycoprotein CD3 zeta chain.
 FT DOMAIN 22 30 Extracellular (Potential).
 FT TRANSMEM 31 51 Potential.
 FT DOMAIN 52 165 Cytoplasmic (Potential).
 FT DOMAIN 69 89 ITAM 1.
 FT DOMAIN 107 128 ITAM 2.
 FT DOMAIN 140 160 ITAM 3.
 FT DISULFID 32 32 Interchain (Potential).
 FT MOD_RSS 154 154 Phosphotyrosine (By similarity).
 SQ SEQUENCE 165 AA; 18773 MW; 3183136130BAA4F5 CRC64;

Query Match 25.1%; Score 589.5; DB 1; Length 165;
 Best Local Similarity 81.6%; Pred. No. 1.6e-35;
 Matches 115; Conservative 9; Mismatches 14; Indels 3; Gaps 2;

QY 305 LEDPKCYLLDGIIFITGVITITALLYLRKRSRSRSETAANLDPNQLYNELNGRREYDV 364
 DB 26 LLDPKCYLLDGIIFITGVITITALLYLRKRSRSGEDVPSQGHQOLYNELNGRREYDV 85
 QY 365 LEKRRADPEMGKQORRRNPOEGVYNALOKDKAAEAYSEIGTK--RRRGKGGHDGLYQ 422
 DB 86 LDRRGFDPEMGKQORRRNPOEGVYNALOKDKAAEAYSEIGTKGGRRGKGGHDGLYQ 144

QY 423 GLSTATKDTYDALHMQTLAPR 443
 DB 145 GLSATATKDTYDALHMQTLPPR 165

RESULT 15
 CD3Z SHEEP STANDARD; PRT: 166 AA.
 AC P293329; DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-DEC-1992 (Rel. 24, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE T-cell surface glycoprotein CD3 zeta chain precursor (T-cell receptor
 DE T3 zeta chain).
 GN Name=CD3Z;
 OS Ovis aries (sheep).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Caprinae; Ovis.
 OC NCBI_TaxID=9940;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=White alpine;
 RX MEDLINE=9313105; PubMed=8420837;
 RA Hein W.R., Tunnicliffe A.;
 RT "Invariant components of the sheep T-cell antigen receptor: cloning of
 RT the CD3 epsilon and Tcr zeta chains.";
 RL Immunogenetics 37:279-284(1993).
 CC -1- FUNCTION: Probable role in assembly and expression of the TCR
 CC complex as well as signal transduction upon antigen triggering.
 CC -1- SUBUNIT: The TCR/CD3 complex of T lymphocytes consists of either a
 CC TCR alpha/beta or TCR gamma/delta heterodimer coexpressed at the
 CC cell surface with the invariant subunits of CD3 labeled gamma,
 CC delta, epsilon, zeta, and eta.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- PTM: Phosphorylated on Tyr residues after T-cell receptor
 CC triggering (By similarity).

CC -1- SIMILARITY: Belongs to the CD3Z/FCER1G family.
 CC -1- SIMILARITY: Contains 3 ITAM domains.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL; Z12968; CAAT7312.1; -
 DR PIR; I46424; I46424.
 DR InterPro; IPR003110; ITAM.
 DR Pfam; PF02189; ITAM; 3.
 DR SMART; SM00077; ITAM; 3.
 KW Phosphorylation; Receptor; Repeat; Signal; T-cell; Transmembrane.
 FT SIGNAL 1 21 By similarity.
 FT CHAIN 22 166 T-cell surface glycoprotein CD3 zeta chain.
 FT DOMAIN 22 30 Extracellular (Potential).
 FT TRANSMEM 31 51 Potential.
 FT DOMAIN 52 166 Cytoplasmic (Potential).
 FT DOMAIN 69 89 ITAM 1.
 FT DOMAIN 108 129 ITAM 2.
 FT DOMAIN 141 161 ITAM 3.
 FT DISULFID 32 32 Interchain (Potential).
 FT MOD_RSS 155 155 Phosphotyrosine (By similarity).
 SQ SEQUENCE 166 AA; 18704 MW; E7D89AD84E58311A CRC64;

Query Match 23.0%; Score 539; DB 1; Length 166;
 Best Local Similarity 74.5%; Pred. No. 8.4e-32;
 Matches 105; Conservative 13; Mismatches 21; Indels 2; Gaps 1;

QY 305 LEDPKCYLLDGIIFITGVITITALLYLRKRSRSRSETAANLDPNQLYNELNGRREYDV 364
 DB 26 LLDPKCYLLDGIIFITGVITITALLYLRKRSRSADAPAYOHQNPVYNELNGRREYDV 85
 QY 365 LEKRRADPEMGKQORRRNPOEGVYNALOKDKAAEAYSEIGTK--GRRRGKGGHDGLYQ 422
 DB 86 LDRRGFDPEMGKQORRRNPOEGVYNELNRDKAAEAYSEIGTKSDNORRRGKGGHDGLYQ 145

QY 423 GLSTATKDTYDALHMQTLAPR 443
 DB 146 GLSATATKDTYDALHMQTLPPR 166

Search completed: April 20, 2005, 07:07:09
 Job time : 64.2157 secs



THIS PAGE BLANK (USPTO)